

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 4, 2001, 18:33:32 ; Search time 6385.92 Seconds
(without alignments)
8930.521 Million cell updates/sec

Title: US-09-117-447-1

Perfect score: 3687

Sequence: 1 atgataagagaaaaagctgt.....ttacattctataataagtaa 3687

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_cm:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_pi:*
12: gb_pi1:*
13: gb_pi2:*
14: gb_pi3:*
15: gb_pi4:*
16: gb_ba1:*
17: gb_ba2:*
18: gb_cm:*
19: gb_ov:*
20: gb_pi1:*
21: gb_pi2:*
22: gb_pi3:*
23: gb_pi4:*
24: gb_pi5:*
25: gb_pi6:*
26: gb_pi7:*
27: gb_pi8:*
28: gb_pi9:*
29: gb_pi10:*
30: gb_pi11:*
31: gb_pi12:*
32: gb_pi13:*
33: gb_pi14:*
34: gb_pi15:*
35: gb_pi16:*
36: gb_pi17:*
37: gb_pi18:*
38: gb_pi19:*
39: gb_pi20:*
40: gb_pi21:*
41: gb_pi22:*
42: gb_pi23:*
43: gb_pi24:*

44: em_ov:*
45: em_pat:*
46: em_pi:*
47: em_pi1:*
48: em_pi2:*
49: em_pi3:*
50: em_pi4:*
51: em_pi5:*
52: em_pi6:*
53: em_pi7:*
54: em_pi8:*
55: em_pi9:*
56: em_pi10:*
57: em_pi11:*
58: em_pi12:*
59: em_pi13:*
60: em_pi14:*
61: em_pi15:*
62: em_pi16:*
63: em_pi17:*
64: em_pi18:*
65: em_pi19:*
66: em_pi20:*
67: em_pi21:*
68: em_pi22:*
69: em_pi23:*
70: em_pi24:*
71: em_pi25:*
72: em_pi26:*
73: em_pi27:*
74: em_pi28:*
75: em_pi29:*
76: em_pi30:*
77: em_pi31:*
78: em_pi32:*
79: em_pi33:*
80: em_pi34:*
81: em_pi35:*
82: em_pi36:*
83: em_pi37:*
84: em_pi38:*
85: em_pi39:*
86: em_pi40:*
87: em_pi41:*
88: em_pi42:*
89: em_pi43:*
90: em_pi44:*
91: em_pi45:*
92: em_pi46:*
93: em_pi47:*
94: em_pi48:*
95: em_pi49:*
96: em_pi50:*
97: em_pi51:*
98: em_pi52:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	3687	100.0	3687	9	A93832 Sequence 1
2	3687	100.0	3687	9	AX000218 Sequence
3	3687	100.0	3687	9	AX033303 Sequence
4	3687	100.0	3687	2	BSSBSA
5	513	13.9	3680	1	AF055578 Bacillus
6	487.6	13.2	3090	2	AF228338 Bacillus
7	77.4	2.1	2069	45	E10125 DNA encodin
8	77.4	2.1	3399	45	E10126 DNA encodin

9	77	2.1	111554	13	AP002460	AE002460 Arabidops
10	73	2.0	5420	5	AF056936	AF056936 Plasmodi
11	71.4	1.9	179553	68	AC024253	AC024253 Homo sapi
12	66	1.8	208006	77	AC087877	AC087877 Mus muscu
13	64.2	1.7	110000	84	PFMA4P1_1	Continuation (2 of
14	64	1.7	3720	59	576368	576368 ORF 5' of E
15	64	1.7	43658	59	HSV3PRGEN	M64409 Herpesviru
16	64	1.7	112930	59	HSGEN	M64346 Herpesviru
17	62.6	1.7	12029	45	AE001373	AE001373 Plasmodi
18	62.4	1.7	1686	45	E08995	E08995 DNA encodin
19	61.6	1.7	423	3	MC049	Z33043 M. capricolu
20	60.2	1.6	4050	6	DDPROKING	Z37981 D. discoiden
21	59.6	1.6	3009	41	AF323001	AF323001 Plasmodi
22	59.6	1.6	16377	5	AF202180	AF202180 Plasmodi
23	59	1.6	1324	96	PERSAR1	X05182 P. falciparu
24	59	1.6	1587	96	PFAL18528	Y18528 Plasmodium
25	59	1.6	4591	9	A00661	A00661 P. falciparu
26	58.8	1.6	4995	96	PFAMESA	M69183 Plasmodium
27	58.4	1.6	195969	75	AC074354	AC074354 Oryza sat
28	57.8	1.6	2196	2	BB078764	U78764 Borrelia bu
29	57.8	1.6	2261	2	BB072996	U72996 Borrelia bu
30	57.8	1.6	2642	5	AF103869	AF103869 Plasmodi
31	57.8	1.6	30750	1	AE001575	AE001575 Borrelia
32	57.8	1.6	30885	1	AE001580	AE001580 Borrelia
33	57.6	1.6	186935	67	AC022322	AC022322 Homo sapi
34	57.4	1.6	4283	2	BB044912	U44912 Borrelia bu
35	56	1.5	145050	70	AC026679	AC026679 Mus muscu
36	55.6	1.5	152409	64	PFMA1P1	AI031744 Plasmodi
37	55.4	1.5	169676	67	AC023334	AC023334 Homo sapi
38	55	1.5	1141	10	AX083744	AX083744 Sequence
39	55	1.5	5173	6	MDG14D5	X53237 D. discoid
40	55	1.5	6797	3	MHMP34GN	X95601 M. hominis
41	55	1.5	10036	56	AF269225	AF269225 Cloning v
42	55	1.5	10036	56	AF269225	AF269225 Cloning v
43	55	1.5	10452	56	AF269238	AF269238 Cloning v
44	55	1.5	11166	56	AF269237	AF269237 Cloning v
45	55	1.5	11335	56	AF270470	AF270470 Cloning v

ALIGNMENTS

RESULT 1					
A93832	3687 bp	DNA			
LOCUS	Sequence 1 from Patent WO9728263.	PAT			
ACCESSION	A93832				
VERSION	A93832.1	GI:6741963			
KEYWORDS					
SOURCE					
ORGANISM	Bacillus steaerothermophilus.				
	Bacillus steaerothermophilus				
	Bacteria: Firmicutes; Bacillus/Clostridium group;				
	Bacillus/Staphylococcus group; Bacillus.				
	1 (bases 1 to 3687)				
	Lubitz/W. and Sleytr/U.				
REFERENCE	1 (bases 1 to 3687)				
AUTHORS	Lubitz/W. and Sleytr/U.				
TITLE	RECOMBINANT EXPRESSION OF S-LAYER PROTEINS				
JOURNAL	Patent: WO 9728263-A 1 07-AUG-1997;				
	LUBITZ WERNER (AT); SLEYTR UWE (AT)				
FEATURES	Location/Qualifiers				
SOURCE	1..3687				
	/organism="Bacillus steaerothermophilus"				
	/strain="PV72"				
	/db_xref="taxon:1422"				
	/clone="SBSA"				
	1..90				
	1..3687				
sig_peptide	/note="unnamed protein product"				
CDS	/codon_start=1				
	/transl_table=1				
	/protein_id="CAB69613.1"				
	/db_xref="GI:6741964"				
	/translation="MDRKKAVKLAATASATAAVANPNASEAATDVATVSSAKAO				
	FKKAVYTSHTVETGEPPINIDVYAEYNKAKRYRDAVALYNKAGGAKKADYLAIDQ				

KEYETVYFKNPKSGSEARVATYIDAVYNATKIDEMROBELNAVOAKDLEKAEQYHHKI
PYEIKRTVYIDRVYKGTTRDLPLSTFEAKQOEIADSLIYITVAMKAREVDQVAKG
NLDRKAAVNDINOYLPKVPVAPTELFEVKKALDEALIDLETTPEVESVAINONKA
VELTAPVNGTKIKLOLSAANEDPVNVNPTVYVDGNIPALNTADVISLSDKRTIT
VDASTPEENNTETIKVYKGIKDKMKEREKEDATPFLKNDVAVIVQRTNNTSVN
LAAGTFEDDTLVVFDKLAPEVNSSNVTTLVEYTKRIPVIAAGSGSTIYITLKE
ALVNGKOYKLAINNKKYLLJGNAEYELFEVFNMSAPVAAPTLGGTTLSTGLT
NMWGLAGVNEAGTYVYGLQFTTFEARKLEESLADAEVLEKESGVASSELKYN
DAKWITVPRKADLEKENTYIOIKIRKKSODGIELGTENKTEFEKEDPLVAPVYS
TSKNDGALPKTEAOEETFEVSESNFTNPTAATVYGGSTITTYQVAVYKGNALSLTAS
DIPASVDAVTGQDSTIKVYKVAANDLENNGIKLVYEGKATVAPKADANNTLATNT
IYPTTEGQDVTAPLVYKVFEGDSLKDADAATVTLNVDAGKFTIYSEELKTSGL
VGGRTVEKELTNNGVDAGTGTFTVSAKFTDANGKVTAAVYTLGLDNDDALRLV
VDKSTDGIADVAGNVIKERDILRYNMRFTVSVKAAADCOGNSAAPSTTAID
TTKSLVFNEDTLAEVYKPEINIVDAAGNVAGVYALDGSSTKFPVTPSOELKAGT
VYSVPTIGVRRKNGNTISKYITTSKRYTSAANTPLSISITADAVYVDSKRTITTESDS
VNPPTITLKRADGTSFTNVTYLVNWNNEKRTIKIYFHKQVTLDEFTQELAVSAPLO
TIDISKVTEITIGSVATDEKPLALVGVGWNCTSYQDAATRLRSVADPVAAPALOF
SEGIDLTNATVYTNITDDKTEVYISKESVDADHDAGATKETLVIINTVPLVLDNSKT
YKIVSGVDAAGNVADITVYIK*

mat_peptide

BASE COUNT

1316 a

660 c

715 g

996 t

Query Match

100.0%; Score 3687; DB 9; Length 3687;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3687; Conservative

0; Mismatches

0; Indels

Gaps 0;

QY	1	atggaatagaataaagctgtgaactagaagaagctgtactgtcagcaagtcatt	60
DB	1	ATGGATAGGAAAAAGCTGTGAACATGACGACGAGTCTATTGACAGAGTCAT	60
QY	61	gtcgcgtgcaaatccaaagctctcgaagcggtctacaagatgtctgaacagtagtaagccaa	120
DB	61	GTGCGTGCAAAACCAACGCTTCTGAAGCGGTACAGAGTGTGCAACAGATGTAAGCCAA	120
QY	121	gcaaaagcaagctcacaagaacactactactcagccatcaagtaagcaagaaactgt	180
DB	121	GCAAAAAGCAGATTTAAAAAAGCATATCTTACAGCCATACGTAAGGAAACTGCT	180
QY	181	gaattcccaaacatcaacgaatgatalatcgtgaatacaacaagaacgaaacgaatccgt	240
DB	181	GAATTCGCCAAACATTAACGATGTATATGCTGAATCAACAAAGCGAAAAACGATACCGT	240
QY	241	gatgcggtgtagcattgtgaataagcaggtggtcggaataaagaagcttacttgaatgat	300
DB	241	GATGCGGTAGCATTTAGTGAATTAAGCAGGTGGCGGAAAAAGACGTTACTTACCTGAT	300
QY	301	ttacaagaagaatagaactacgttttccaagaacacccttaactctggtgaagctcgt	360
DB	301	TTACAAAAGAAATATGAANAATCTACGTTTCAAGCAAAACCCCTTAATCTGGCAACCTCGT	360
QY	361	gtagaacttacaatcatgatgtcttaacaactatgcaacaanaatagacgaatgctgcgaag	420
DB	361	GTAGCAACTTACATCGATGAGCTTACCAACTATGCAAAAATTTAGAGCAAAATGCGCAAGAG	420
QY	421	ctgaagcgctgctgtccaagcaaaaatttagaagaagaagaacaatatactatcaacaatt	480
DB	421	CTAGAGGCTGCTGTTCAAGCAAAAGATTTAGAAAAGAGAAACAATTAATTCACAAATTT	480
QY	481	cccttgaataataaacgcgcacagctcatcttagatcgcgtatcgttatacgttaaaacactcgt	540
DB	481	CCTTATGAAATTAATAACCTCGCACACTCATTTTATGATTCGGTATATGTTAAACCAACTCGT	540
QY	541	gattactactcgctacatttaaagcaaaagcacaagaactctgcgacagcttaattat	600
DB	541	GATTACTTCCGCTYACATTTTAAAGCAAAAGCAAAAGCACTTCGCGACAGCTTAATTTAT	600
QY	601	gattatcagctgtgaataaagcgcggaagtagacaagaagcgtgtgtgaagcgagcaattta	660
DB	601	GATTATTAACGTTGCAATGAAGCGCGCAAGATPACAAAGCGCTGTGAAGGACAGCAATTTTA	660

Db	1741	GAATAATACAACTCTCAAAATCAAAAATTTAAAAAGGCTTGAAAGTCGATAAAGCTATATGAA	1800
Oy	1801	ttgaagcattgttlaacgcggaanaacaatagagttcaaaactaaacttaactctgcttaca	1866
Db	1801	TTAGGCACTGTTTAACGAGAAAACATATGATGTTCAAAACCTAGACTTTAACTGCTCCTACA	1866
Oy	1861	gttatttagcgtlaacgctctaaanaatlygcgcgcgtlgtataaaglaaclygaagctcaagaa	1920
Db	1861	GTTATTATACCGTAACGTCCTAAAAATGGCGAGCGTGGATTTAAATTAACCTAGACTCAAGAA	1920
Oy	1921	tttactgtggaagttcctcaagagaatttaatafacattaaigtctacaacggtttccggttagc	1986
Db	1921	TTTTACTGTGGAAGTTCTCAGAGCAATTTAAATACATTTAAATCTCAACACGTTTCGGGTAGC	1986
Oy	1981	acaatcacatiacggttcaagttgctgtgtatgaagccgggtgcgaactatctgcttaca	2044
Db	1981	ACAATCACATACGGTCAAGTTGCTGTGTGTAAAGCGGGTCAACACTTAATCTGCTCTTACA	2044
Oy	2041	gcaagttgacatcatctccagctagttgttgaagccggttactggtlccaagaatgaacatacaaa	2100
Db	2041	GCAAGTGCACATCATTCCACGAGTGTGAAGCGGTTACTGTCGAAGATGGAACATACAAA	2100
Oy	2101	gtgaaagttgctgtcttaaccaattagaacgctlaaccaaggggtacaaattagtagtgtccgtt	2166
Db	2101	GTTGAAGTTGCTGCTMAACCAATTGAGACGTAACCAAGGGTCAAAATTAAGTAGTGTCCGT	2166
Oy	2161	aaagttgnaaagctcctcgtttaagaatgctgcgaanaatgcgaactctttagcaactaactat	2222
Db	2161	AAAGGTGCACACGTCCTCGTTAAAGATGCTGCAAAATGCAAAATCTTTACCACTAACTAT	2222
Oy	2221	atctatcaattacaacttgaagtgcaagacgtlaacagccaccaacggtltacaagaatattc	2286
Db	2221	ATCTATCAATTTACACTGGAAGGTCAAGAGCTATACGACACCAACGGTTTCAAAAAGTATTC	2286
Oy	2281	aaagttgattctcttaaaagacgctgaltgcagttacttaacttaacgaagltgalgcaggt	2340
Db	2281	AAAGTGTATCTTTAAAAAGACGGTGAAGCTGATTACTACACTTACGAACGTTGATGACGCT	2340
Oy	2341	caaaaattcactatccaatlltagcggaagaatttaaaactctagtggttctttagtgggt	2400
Db	2341	CAAAAATTCACATTCACAAATTTAGCGGAAGATTAAAACTTCAATGCTGTTAGTGGCT	2400
Oy	2401	gccaagtaactgtgcggaanaatlaacaaacacggatlggtlagaatgcctgtactcgtgaaca	2466
Db	2401	GGCAAAAGTAACCTGTCGAGAAATTAACAACAACGGATGGGTAGATGCTTGCTACTGGAACA	2466
Oy	2461	actgtatacgttgcctcctaagacacagatgycnaatlygtlaaaatgaacagcttgcctgttaca	2520
Db	2461	ACTGTATCAGTTGGTCCCTAAGACAGATGCAAAATGATTAAGTAAACACTGCTGTGGTTACA	2520
Oy	2521	tttaactgctcttgaacaataacgacaaagaatgcgaaatlgcgtctggtagtagataagtc	2586
Db	2521	TTTAATGCTCTTGGCAATTAACGACAAAGATGCCAAATTTGGTCTGTAGTAAATGAGTCT	2586
Oy	2581	tctaactgtagaattgctgtatgtagtcgtgtgaatgtaatttaagaaaaagaattttaatt	2640
Db	2581	TCTACTGTATGGAATTTGCTGATGATGCTGATTAAGTAAATTAAGAAAAAATATTTTAAAT	2640
Oy	2641	cgttacaacagcttgagagacacactgttagcttccgttgaagcttgcctgtctacaaagaatggt	2700
Db	2641	CGTTTACAAACACTGGACACACACAGTATGCTTCTGTGAAGCTGCTGCTGACAAAGTGGT	2700
Oy	2701	caaaaagctctgtctgcatatcccaacaagcactgcgaattgatatacaactaagaacttata	2766
Db	2701	CAAAAACGCTTCTCTGATTCCTCCACACACACACGACATGTGATTAACAATCAAGAGCTTATTTA	2766
Oy	2761	gttaattcaatgaagaattttagcggaaatgaaaaccgtgaagaactcgttgttaaagat	2820
Db	2761	GTTCAATTCATGTAACCTGATTTAAGCGGAATTTAAACCTGAGAACATCGTGTTTAAAGAT	2820
Oy	2821	gcacaggttaatgcggtagatggttactgtlaacagcaattgaagcgtttctacaataaattc	2886

DB 2821 GCAGCAGTAATCGGTAGCTGCTACTGTAACAGCATTAAGCGGTCTCTCAATAAATTT 2880
QY 2881 gatttaccctccctcctaagaattaaagctggtacagtttactctctgtaacaattgcggt 2940
DB 2881 gatttaccctccctcctaagaattaaagctggtacagtttactctctgtaacaattgcggt 2940
QY 2941 ggtgaagataagtagtgtaacacaatctctaataatactactctgctcaagactgctc 3000
DB 2941 ggtgaagataagtagtgtaacacaatctctaataatactactctgctcaagactgctc 3000
QY 3001 gctgaatccaagcttatctcctaactcagaactgctgacggtgcagttacgttgacgct 3060
DB 3001 gctgaatccaagcttatctcctaactcagaactgctgacggtgcagttacgttgacgct 3060
QY 3061 aaaaacattacaattgaattcaggagttcagtttccaaaccacaactcactcttaagaag 3120
DB 3061 aaaaacattacaattgaattcaggagttcagtttccaaaccacaactcactcttaagaag 3120
QY 3121 gctgacggaactcattcttaactaactcactctgtaataatgtaataatgtaataatgta 3180
DB 3121 gctgacggaactcattcttaactaactcactctgtaataatgtaataatgtaataatgta 3180
QY 3181 tacaataatgtaattccacaagaagtgtaacacttgacgagttcactcaataatgagttaga 3240
DB 3181 tacaataatgtaattccacaagaagtgtaacacttgacgagttcactcaataatgagttaga 3240
QY 3241 gtttcaaaaagatttcaaaactgctgactgattgagatgacaaattgacattcactcagag 3300
DB 3241 gtttcaaaaagatttcaaaactgctgactgattgagatgacaaattgacattcactcagag 3300
QY 3301 tctgtgtcactgacgaagtaaaacctgctcagtagtcgttggttcattggaattggaaca 3360
DB 3301 tctgtgtcactgacgaagtaaaacctgctcagtagtcgttggttcattggaattggaaca 3360
QY 3361 agccttaccagaagatgctcagaacagacagactcagctgctgactgactgctgctgctg 3420
DB 3361 agccttaccagaagatgctcagaacagacagactcagctgctgactgactgctgctgctg 3420
QY 3421 ccagttgccttcaactctcagaaggtatcgaattcaacgaatgcaactgtgaacgttaaca 3480
DB 3421 ccagttgccttcaactctcagaaggtatcgaattcaacgaatgcaactgtgaacgttaaca 3480
QY 3481 aatattactcgatgataaaactgtggaagtattcacaagaagagtgtagaagcagacat 3540
DB 3481 aatattactcgatgataaaactgtggaagtattcacaagaagagtgtagaagcagacat 3540
QY 3541 gatgcaggtgactcagaagagacattgaattcaacagactgactccttactgactgact 3600
DB 3541 gatgcaggtgactcagaagagacattgaattcaacagactgactccttactgactgact 3600
QY 3601 aacagcaagactataaagattgttgtaagtgaggttaagaatgacgacgtaattgttga 3660
DB 3601 aacagcaagactataaagattgttgtaagtgaggttaagaatgacgacgtaattgttga 3660
QY 3661 gatactattacattctatataatgaata 3687
DB 3661 gatactattacattctatataatgaata 3687
RESULT 2
AX000218 3687 bp DNA PAT 10-MAR-2000
LOCUS AX000218 Sequence 1 from Patent WO9006567.
ACCESSION AX000218
VERSION AX000218.1 GI:7240682
KEYWORDS
SOURCE
ORGANISM
Bacillus stearothermophilus.
Bacillus stearothermophilus
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
REFERENCE
1 (bases 1 to 3687)
Lubitz,W. and Resch,S.
AUTHORS
TITLE
SECRETION OF CARRIER-BONDED PROTEINS INTO THE PERIPLASMA AND THE

EXTRACELLULAR SPACE
Patent: WO 9006567-A 1 11-FEB-1999;
JOURNAL LUBITZ WERNER (AT); RESCH STEPHANIE (DE)
FEATURES
source
1..3687
/organism="Bacillus stearothermophilus"
/strain="pv72"
/db_xref="taxon:1422"
/clone="SBSA"
1..90
1..3687
sig_peptide
CDS
1..3687
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CA877069.1"
/db_xref="GI:7240683"
/translation="MDRRKAVKLTATASIAASAFYAAPNPNSAATDVATVVSQAKQ
FKKAVYTSHTVETGERPNINDVYAEKNKKRKYRDVALVNRKAGKRDVATVVSQAKQ
KEYETVYFKNRPSGEARVATYIDAYNATKLDKDEKLEDAVQAKDLEKAEQYHKI
PYEIKRTVILDRYVKKTRDLRSTFKAKOEKRLDSLIYITVAMKAREVODAVKAG
NLDRAKAAVDIOINQYLPKVTDAFKTELFEVAKKALDADEALTPKVESVSAINTONKA
VELTRPVNGLKLDLSAANEDVYNVNTVATIKYDGNIPRALNTADVSLSTDGRTIT
VDASTPEENNTETKVVVNGIKDKNGKEKEDAFTEFKLNDVAVTQVETNTNTSVN
LAAGFPDIDPLTVVFDKILAEPEVNSSNVTITVEGKRILPIVASTGSITITTLKE
ALVNGKQYKLAINNKKTLGYNAEYELVFTNMSAPVATPVLVEKSGTVASSELKLT
NMVGLAGVNEAGTYYPGLQFTTTFEAKLESTLADNPLVEKSGTVASSELKLT
DAKWTLLVKNADLEKNTIYQIKKKGLSDGIELGVNERTYEERKTDPLAPYISV
TSKNGDALKYTEQEFKVESENLPNATTVSGSTITTVGOVAVVNRKAGNLSALTAS
DIIPASVAVYGGDSYTVKRYANDLENQGYKLVFRRKGTATPKDAAANNTLATNY
IYTFTEGSDVATPVTIKVFKGDSLKDADAVTTLNVDAQGFITIOSEELKTSGL
VGGKSTVERKLNNGVNDAGSTTVSVARTKANGKVAAVVTGLDNNDADAKRLV
VDSKSTDEIADVAGNVIKEDKILIRYNSRHTVSVKRAADQDNASAPPTSTAIID
TTKSLVFNEDTDLAEVKEPNIIVVDAGNNAVAGVTLDSTNKFVPTPOEKLKAGT
VYSVITDGVROKGNITISKYITSEKTVANPLTSSISIAQDADVNRKKTITIEKSDS
VNPNTITLKKADGNSFTNVTLVNANNENKRTKIVFRRKGTITDEFTQVLEAVSKPQNS
TDIDSKVTFITGSVATIDEVKRALVGVGSMKNTSTYQDPAATRLSVADFAVAPLALQF
SEGIDLVNATVTNITDQKTEVEVSKESVDADHDAGATKELTIVINTVPLVNLQKT
YKIVSGVKDAAGNVADRTITFEYIK"
mat_peptide
BASE COUNT 1316 a 660 c 715 g 996 t
ORIGIN
Query Match 100.0%; Score 3687; DB 9; Length 3687;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 atggaataggaataaagctgtgaaactagcaacagaagtgcataatgcaattc 60
DB 1 atggaataggaataaagctgtgaaactagcaacagaagtgcataatgcaattc 60
QY 61 gtcgctgcaaatccaaagctctcgaagcgcttacagatgtagcaacagtagtaagccaa 120
DB 61 gtcgctgcaaatccaaagctctcgaagcgcttacagatgtagcaacagtagtaagccaa 120
QY 121 gcaaaagcagctcaaaaagcatatactatactcagcgcatacagtaacggaactgct 180
DB 121 gcaaaagcagctcaaaaagcatatactatactcagcgcatacagtaacggaactgct 180
QY 181 gaattcccaaacattaaagatgatatgctgtaatacaacaagaagcaaaaacgataccgt 240
DB 181 gaattcccaaacattaaagatgatatgctgtaatacaacaagaagcaaaaacgataccgt 240
QY 241 gatgcggtgacattgtagtaataagcagtgacggaataaagaagcttactgactgact 300
DB 241 gatgcggtgacattgtagtaataagcagtgacggaataaagaagcttactgactgact 300
QY 301 ttacaanaagatagaactacgctttccaagaacaaaccctaactctgcggaagctcgt 360
DB 301 ttacaanaagatagaactacgctttccaagaacaaaccctaactctgcggaagctcgt 360
QY 361 gtagcaacttaccatgactgactcacaactatgcaacaaaatagacgaatgctgcgaagag 420
DB 361 gtagcaacttaccatgactgactcacaactatgcaacaaaatagacgaatgctgcgaagag 420

```
Db 361 GTAGCACTTACATCGATGCTTACAACTATGCAACAAATTTAGACGAATGCGCCAAAG 420
Oy 421 cttagagcgtgcgttcaagaacaaagtttagaaaaagcgaacaaatctatccaaatt 480
Db 421 CTAGAGGCTGCTGTTCAACGAAAAAGATTGAAAAAGCAAGAACAAATCATCTCAAAATT 480
Oy 481 ccttaagaattaaaactcgcaacagctcatctttagatcgctatagtgtlaaacaactcgt 540
Db 481 CCTTATGAATTTAAACCTGCGACAGTCATTTTATGATCGGATATATGTATAAACAACTCGT 540
Oy 541 gatttacttcgctctacattttaaagcaaaagcaaaagccttcgcgacagcttaattat 600
Db 541 GATTTACTTCGCTCATATTTTAAAGCAAAAGCACAAGAACTTGCGACAGCTTAATTTAT 600
Oy 601 gatattaccgtttgcgaatgaaagcgcggaagtacaaagcgtgtgtgaagcaggcaattta 660
Db 601 GATATTATTCGGTTGCAATGGAAGCGCGGAAGTACAAGACGCTGTGAAACCAAGCAATTTTA 660
Oy 661 gacaaagcctaaagctgtgttgatcaatccaatccaattaccacaaagttaacagatgtc 720
Db 661 GACAAAGCTAAAGCTGCTGTTGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
Oy 721 ttcaaaactgaaactaaacagaagttagcgaanaaagcattagatgcagatgaagctgcctt 780
Db 721 TTCAAAACTGAAGCTAACAGAGTAGCGAAAAAGCATTAGATGCAGATGAGCTGCGCTT 780
Oy 781 actccaaaagtttgaaagtgtgaagtgtgcatcttaacaccctcaaaagcgtgttgaaataca 840
Db 781 ACTCCAAAAGTTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 840
Oy 841 gcaagtaaccagtgaaacgaacactaaatcacacttcagctgtgcgcaatgaaataga 900
Db 841 GGAGTACCGAGTGAACGGAACACATAAATTTACAACTTTCAGCTGCTGCAATGGAAGTACA 900
Oy 901 gtaaacgttaatactgttactgtatctataaagtggagcgtaaacattccatttgccttaat 960
Db 901 GTAAACGTAATACTGATGATCTCTTAAAGTGAGCGATGACATTCATTTGCGCTTAAT 960
Oy 961 agcgaagatggttcttactctacagcggaaaaactatcacgtgtgagttgttcaactcca 1020
Db 961 AGCGGAGATGTTCTTTATCTACAGACGGAATAAATCATCTGATGCTTCAACTCCA 1020
Oy 1021 ttccgaataatactagagatataaagttagttagttaaagtttaagaacaaatgtgcaaa 1080
Db 1021 TTCGAAATAATATACGAGATATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1080
Oy 1081 gaatttaagaagaatgcattcaacttcgaagcttgcgaatgagctgttacttcaagtgt 1140
Db 1081 GAATTTAAAGAAAGATGCAATTCATTCAAGCTTCGAAATGATGCTGTACTCAAGTG 1140
Oy 1141 ttgtgaactaatgttaacaaacaaacactctgtfaaacttagcagcaggtacttctcgact 1200
Db 1141 TTTGGAACTAATGTAAACAAACACACTTCTGTAACTTAGCGACAGGTACTTTCGACACT 1200
Oy 1201 gacgaactttaaagtagtatttgaataagtgtttagcaactgaaactgttaaacaagctgtg 1260
Db 1201 GACGATACTTTAAAGTAGTATTGTATAGTGTGTAGCACTGAAACTGTAAACACACTGCG 1260
Oy 1261 aacgttactatcaacagatgttgaactgtgaaaaacgcatacccgtaattgcactactct 1320
Db 1261 AACGTTACTAATTAAGATGTTGAAACGTGAAACGCATTCAGTAATTCATCTACTTCTT 1320
Oy 1321 ggtctacattactatctacgtcttaaagaagcgtttagtaacgtgttaacaataataaact 1380
Db 1321 GGTCTTACAACTATTAATGCTTTAAAGAGAGCTTAGTAACTGTAAACAAATTAATTAACCTT 1380
Oy 1381 gctatcaaatatgtttaaacaatlaactgtgttacaatgcaagaagcttacagatagtgctc 1440
Db 1381 GCTATCAATTAATGTTAAACATTAACTGTTACATGCTTACAGAGCTTACAGAGTTAGTGTTC 1440
Oy 1441 actgaaaagcgtatcagcccaactgtgtgtacacgctccctcaacttaactttagtgtaaac 1500
Db 1441 ACTGAAAAGCGATCGACACCAACTGTGTACCGCTCTTAATTAATTAATTAATTAATTAATTA 1500
Oy 1501 ttaactacgtgtctcttacaacaaatgtttgggttaaatgtgctgtgtgtgtgaatga 1560
Db 1501 TTATCTACTGCTGCTCTTCAACAAATGTTTGGGGTAAATGTGCTGTGTGTGAATGA 1560
Oy 1561 gctggaacttatatccctgtcttcaatccaacaaacgcttctactaagttagagcaaa 1620
Db 1561 GCTGGAACCTTATATCTGCTGTCTTCAATTCACAAACAGCTTCTCTACTAAGTTAGACGA 1620
Oy 1621 tctacttagctgataacttgttatttagttgaaagaagactcgtgtaacgttgttctct 1680
Db 1621 TCTACTTAGCTGATTAACCTTTGATTAAGTTGAAGAAAGAAATCTGTACAGTTGTTGCTTCT 1680
Oy 1681 gaactaaataataatgcagaagcttaaaatgtgttaactttagtgcacaaagcgaacttaa 1740
Db 1681 GAACATAAATATATATCAGACGCTAAATGCTTAATGCTTAATGCTTAATGCTTAATGCTTAAT 1740
Oy 1741 gaaatatacaatctatcaaatcaaatcaaatcaaatcaaatcaaatcaaatcaaatcaaat 1800
Db 1741 GAAATATCAATCTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1800
Oy 1801 tttagcactgtttaaagcaagaacataatgatttcaaaaactcaagacttaactgtctctaca 1860
Db 1801 TTAGCACTGTTTAAAGCAAGAAACATATGAGTTCAAAACTCAAGACTTAATCTCTCTTACA 1860
Oy 1861 gttatttagcgttaacgtcttaaaaatggcgagcgtgtgtgtttaaagttaactgaagctcaaga 1920
Db 1861 GTTATTAGCGTAAAGCTTAAATGCGTAAATGCGTAAATGCGTAAATGCGTAAATGCGTAAAT 1920
Oy 1921 tttaactgtgaagttcttcaagaagatttaaatatacttaattgctcaacacgttttcgggttagc 1980
Db 1921 TTTACTGTGAAGTCTTCAAGAAATTTAAATACATTTTAATGCTTCAACCGTTTGGGTAGC 1980
Oy 1981 acaatcaatacogtcaagtttgcgttagttaaagcgggtgtgcaaatctatctgtcttaca 2040
Db 1981 ACATATCAATACGCTCAAGTCAAGTGTGTAGTAAAGCGGTCGCAAACTTAATCTGCTTTACA 2040
Oy 2041 gcaagagacatacttccagcttagtgttgaagcgggtgtgtgtgtgtgtgtgtgtgtgtgtgt 2100
Db 2041 GCAAGTGAATCATCTTCCAGTCAATGTTTGAAGCGGTTACTGTGTGAATGGAATGGAATGGAAT 2100
Oy 2101 gtgaagaattgcgtcttaacaaattagaagcttaacaaaggttacaatcttagttagtgcgtt 2160
Db 2101 GTGAAGTGTGCTCTTAACCAATTAGAAGCTTAACCAAGGCTACAAATTAAGTGTGCTT 2160
Oy 2161 aaagtgtcaacagctccgtgttlaaagaatgctgcgaatgcaatactttagcaactaat 2220
Db 2161 AAAGTGTCAACAGCTCTCTGTTAAAGATGCTGCAAAATGCAAAATGCAAAATGCAAAATGCAAA 2220
Oy 2221 atctatatactttaaacttgaagttcaagaagcttaacagcgaacaaaggttacaagaatctt 2280
Db 2221 ATCTATATCAATTTAACAATGAAGGTCAAGACGTACAGCAACGAGTTACAAAAGTATTC 2280
Oy 2281 aaagtgtacttctttaaagaagcgtgtgtagtactacacttcaagaagcttgaatgcaggt 2340
Db 2281 AAAGTGAATCTTTTAAAGACGCTGATGCAAGTACTATCACTTAAGCAAGTGTGATGCAAGT 2340
Oy 2341 caaaaattcaactatccaatttagcgaagaattaaaaacttctagtggttcttagtgggt 2400
Db 2341 CAAAAATTCACATATCCAAATTTAGCGAAGATTTAAAAACTTCACTGATGCTTTAGTGGGT 2400
Oy 2401 ggaagaatgaactgtgagaagaattaaacaaacaaaggtgtggtagatgctgttagtgaaga 2460
Db 2401 GCGAAAGTAACTGTGAGGAATTAACAAACAAAGGATGGGTAGATGCTGTACTGGAAGA 2460
Oy 2461 actgtatcagttgtcttcaagacagatgcgaatggttaaaatgaacagctgtgtgtgttaca 2520
Db 2461 ACTGTATCAGTTGCTCTCTTAAGACAGATGCAAAATGTAAGTAACAGCTCTGTGATTACA 2520
Oy 2521 ttaactgtgtcttgacaataacagcaagaatgtcgaaattgtgtgtgtgtgtgtgtgtgtgtgt 2580
Db 2521 TTAACTGTGCTTGACAAATTAACGCAAAAGATGCGAAATGTGCTGTGTGTGTGTGTGTGTGT 2580
```

QY	2581	tcctctgataagaaatgctgtgtgtgtagctgcggtgaatgtaataaaggaaaaagatattttaa	2640
Db	2581	TCCTACTGTATGGAATTTGGTGTGATGTAGCGTGGTAATGTAAATTAAGGAAAGATATTTTAAT	2640
QY	2641	cgattcaacaacacttbgagaacacacttgaagctctcgtgtaaagctgcgtctgacaagaatggt	2700
Db	2641	CGTTACAAACACACTGTGAGACACACACTGTAGCTTCTGTGTAAAGCTGCTGTACAAACATCGT	2700
QY	2701	caaaaagctctcgtgcattccacaagaacactgtcaattgatalacaactaagaagcttata	2760
Db	2701	CAAAAAGCTTCTGCTGCCATTCGCCAACAGACACTGCATTTGATATCAATAGAGCTTATTA	2760
QY	2761	gttgaattccaatgaagaatgtatttagcgggaagttaaaccggaacaatcgtgtgttaaat	2820
Db	2761	GTTGATTCATATGAACCTGATTTTAGCGGAAGTTAAACCTGGAACACATCGTGTGTTAAACAT	2820
QY	2821	gcagcaaggtlaatgctgctgagctggtgtactgtlacaagcaattagaagcttctacaataaatt	2880
Db	2821	GCAGCAGGTATATGGCGGTATAGCTGGTACTGTATACAGCATTTAGACGGTTCTACAAATTAAT	2880
QY	2881	gtattcactccattccaagaattaaagaactgtgtacagtttactctgtaaacaattgaagct	2940
Db	2881	GTATTCACCTCCATCTCAAGAAATTAAGAGCTGTACAGTTACTCTGTAAACATTTAGCCGT	2940
QY	2941	gtgagagataagtagaggtlaaaccaactctcraaataacttaacttcgtttaaagacttact	3000
Db	2941	GTGAGAGTAAATGATAGGTATACACAACTCTCAATTAATTAATTAATTAATTAATTAATTAAT	3000
QY	3001	gcgaatccaagctatcttccaatcaagattgtcgaagctgtgaagctttaaagcttccatct	3060
Db	3001	GCGAATCCAAGCTTATCTTCATTCAGCATTTGCTGACGGTGTCCAGTTTACCTTACCGTTC	3060
QY	3061	aaaacaattacaattgaattcgaagcttcagttccaaaccacaacaactcacttctaagaag	3120
Db	3061	AAAACAATTACAATTGAATTCAGAGATTCAGTTCCTCAACCCCAACATCACTTTAAGAG	3120
QY	3121	gctgaacgaacttacttacttaattacacttttagtaatgtaataatgtaataataaaca	3180
Db	3121	GCTGACGGAACCTTCAATTTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	3180
QY	3181	tacaacaattgattcacaagaaggtgtaaacactgtgaacgagtttaccacaataagaattaga	3240
Db	3181	TACAACAATTGATTCTCCCAAAAGGTGTACACTTGACAGAGTTTACTCATATATGATTTACCA	3240
QY	3241	gtttcaaaaagaatttccaacttggtagtcatgatgatgatgataagcaagttacattccaacagt	3300
Db	3241	GTTTCAAAAAGATTTTCAAACTGGTAGTGAATATGATAGCAAGATTACATTCACAGAGT	3300
QY	3301	tctgttgcatactgcgaagaataaaacctgctctgaatgagcggttgcataatggaatggaaca	3360
Db	3301	TCTGTGCTACTGTGACGAAGTAACCACTGCTGTGTAGCGGTTGTTCAATGATATGGAACA	3360
QY	3361	agctatactgaagtgtagtgaagaacaacagacttcggtctgtatgcgaacttcgttgcgag	3420
Db	3361	AGCTATACTCAGATGCTGTGACGACACAGCACTTCGCTGTGTAGCTACTTCGTTCCGAG	3420
QY	3421	ccagttgcacctcaaatctcagaaggtatcatgatattcaagcaattgcaactgtgtgacgtataca	3480
Db	3421	CCAGTTCGCCCTTCAATTCCTCAGAAAGGTATGATTTAAGCAATGTATGACATTAACA	3480
QY	3481	aataattactgataaacaactgttgaagttatttccaagaagagagtgtgaacgcagacat	3540
Db	3481	AATATTAATCTGATGATTAACCTGTTGAAGTTATTTCAAAAGAGAGTGTAGACGACACAT	3540
QY	3541	gatacgagtgtagtactaaggaggaacttagtaataaacaagttaccctttagtactgtat	3600
Db	3541	GATCGAGTGTCTACTTAAGGAGACATTATGTAATTAACACAGTTATCTCTTTAGTACTTGAT	3600
QY	3601	aacagcaagacttataagattgtgttgaagtgaggtttaaagatgtgaagcagaggttaatgttca	3660
Db	3601	AACGCAAGACTTATAAGATTGTTGTAAGTGAAGTTAAAGATGCCAGAGTATATGTTGCA	3660
QY	3661	gatactattacattctatattaaagtaa687	

[illegible]

```
QY 181 gaatcccaacattaaagatgtatagtctgtaatacacaagaacgaaaaacatacct 240
   |||||
Db 181 GAATTCCTCCAAACATTACATGATATATGCTGAATACAAACAAAGCAAAAACGATACCGT 240
QY 241 gatgcygtaagcatltagtgaataaagaagctgycgcgaaaaaagacgttactaagctgat 300
   |||||
Db 241 GATGCGGTAGCATTAAGTAATAAGCAGGTGGCGCAAAAAGACGCTTACTTAGTGAT 300
QY 301 ttacaaaaaagaatagaacattacgttttccaagaacaaaccttaactctggcgaagctgt 360
   |||||
Db 301 TTACAAAAAAGAAATATGAACCTTACGTTTCAAGCAACCACTTAAATCTGGCGAAGCTCGT 360
QY 361 gtagcaactacatcgatcttacaactatgcacaaatitagaagaaatgycgcaagag 420
   |||||
Db 361 GTAGCAACTTACATGCATGCTTACACATATGCACAAAATTTAGACAAATGGCCCAAGAG 420
QY 421 ctgagagctgctgttcaagcaaaaagattagaaaaagcgaacaaactatcaaaaatt 480
   |||||
Db 421 CTAGAGGCTGCTGTTCACAAACAAAGATTAGAAAAAGCAGAACATATCATCAAAATT 480
QY 481 cctttaaataaataaactgcacagctcatctttagatcggtatagtgtaaaaaactcgt 540
   |||||
Db 481 CCTTATGAATAATTAACCTGCACAGTCATTTAGATCGGTATATGGTAAACAACTCGT 540
QY 541 gatttactcgtctacatttaagaacaaagcacaagaacttcgcgacagcttaattat 600
   |||||
Db 541 GATTTACTTCTCGCTCTACATTTAAAGCAAAAGCACAAGAACTTCGCGACAGCTTAAATTTAT 600
QY 601 gatattacgctgtgcaatgaaagcgcggaaglaacaaagcgtgtgaaagcaagcaattta 660
   |||||
Db 601 GATATTACCGTGTGCAATGAAAGCGCGCAAGTACAAAGACGCTGTGAAGAGCGCAATTTA 660
QY 661 gacaaagcctaaagcgtgtgtgtgtatcaaatcaatcaacttaccaaaagtaagaatgct 720
   |||||
Db 661 GACAAAGCTTAAAGCTGTGTGTGATCAAAATCAATCAATCTTACCAAAAGTAAAGATGCT 720
QY 721 ttcaaaactgaactaacagaaagtagcgaaaaaagcattagatgcagatgaagctgctt 780
   |||||
Db 721 TTCAAAACCTGAACCTAACAGATAGCGAAAAAGCAATTTAGATGCACATGACAGCTGCGCTT 780
QY 781 actccaaaagctggaagtglaagtgcgatlaaacactcaaaaacaaagctgttgaattaca 840
   |||||
Db 781 ACTCCAAAAGTTGAAAAGTGAAGTGAAGTGAATTAACACTCAAAACAAAGCTGTGTAATTAA 840
QY 841 ggaagtaacggtgaagcgaacactaaataataactttcgcgtgtgcgaatgaagtlaca 900
   |||||
Db 841 GGAAGTACCGGTGAAGCGAACTAAATTAACAATTTACAGCTGCTGCAAAATGAAGATACA 900
QY 901 gtaaacgttaatactgtatgtaactaataagtggaagcgtlaacaattcatttgccttaat 960
   |||||
Db 901 GTAAACGTAAATACCTGTACGTATCTATTAAGTGAAGCGGTACACTTCCATTGGCCCTTAAT 960
QY 961 acggaagatgttcttattatctacagaacggaaaaaatactacactgtgtgaatgttcaactca 1020
   |||||
Db 961 ACGGCAAGATGTTCTTATCTATCTACAGACGGAATAAACTATCATCTGTGATGCTTCAACTCA 1020
QY 1021 ttccgaaaaataacagagatataaagttagttaaagtgattaaagaacaaaaatgycaaa 1080
   |||||
Db 1021 TTTCGAAAAAATATACGAGATTAAGATAGTTAAAGGTATTAAACAAAAAATGGCAAA 1080
QY 1081 gaattaaagaagatgattcacttcaagcttcgaagctgaagatgagctgttagtctaagtg 1140
   |||||
Db 1081 GAATTTAAAGAAAGATGCAATTCATTCAAGCTTTCGAATGATGCTGTAGTTACTCAAGTG 1140
QY 1141 ttgtgaactaatgtlaacaaacaaacttctgtlaaactltagcagcaggtlaacttgcgaact 1200
   |||||
Db 1141 TTGTGAACCTAATGTAAACAACAACACTCTGTAAACTTAGCAGCAGATCTTTCGACACT 1200
QY 1201 gacgatactttaacagtagtatttgaataagtggttagcaccgtgaacgtttaaagcgtcg 1260
   |||||
Db 1201 GACGATACTTTTAAACAGTATATTGTGATTAAGTGTTAGCACTGGAACCTGTAAACACACTCG 1260
QY 1261 aacgttactattacaagatgttgaaacgtgaaaaacgcatcccgtaatttgatctacttct 1320
   |||||
   ..
Db 1261 AACGTTACATTTACATATGTTGAAACTGGAACAAACGATTCACATATTCGATCTACTTCT 1320
QY 1321 ggttctacaattactatctacgtttaaagaagcgtltagtaactgtgtaaacaataataact 1380
   |||||
Db 1321 GGTTCACAAATTAATTAAGTATAGTTAAAGAAAGCTGTAGTAACCTGTATTAACATTAACAT 1380
QY 1381 gctatcaataatgtttaaacattaaactgtgttacaatgacgaagcttaacaggttaagtgctc 1440
   |||||
Db 1381 GCTATCAATTAATGTTTAAACATTAACTGGTTACAAATGCAAGAACCTTACAGATTAAGTTC 1440
QY 1441 acgcgaacgcacacagcaccacactgtgtgtacacgcctcactacttaagtgtgacaact 1500
   |||||
Db 1441 ACTGCAAAACGACATCAGACACCAACTGTGCTACCGCTCTTAATCTTATAGTGGTACAACT 1500
QY 1501 ttatctactgtgttcttcaacaacaaatgttgggttaaatltggtgtgtgtgtgaatga 1560
   |||||
Db 1501 TTATCTACTGCTTCTCTTACAAACAAATGTTGGGCTTAATGTGCTGTGTGTGAATGA 1560
QY 1561 gctggaacttatatactcgtgttcaattcaacaacaaatgtgtcttaagttaagcga 1620
   |||||
Db 1561 GCTGGAACCTTATATCTGCTGTTCATTAATTCACAAACAGTTTGTCTAAGTTAGACGA 1620
QY 1621 tctactttagctgataactltgtattagttgaaaaaagaatctgttacaagttgtgtctct 1680
   |||||
Db 1621 TCTACTTTACCTATTAACCTTTGATTAAGTTGAAAAAAGATCTGTGTACAGTTGTGCTTCT 1680
QY 1681 gaaactaaaatataatgacagcgtlaaaatltgtaactttagtgcacaaagcgcaccta 1740
   |||||
Db 1681 GAACCTAAATATTAATGACAGCGCTAAATATGTTAACTTTAGTGCAAAAGCGGACCTTAA 1740
QY 1741 gaaaatacaatctataatcaaatltaaaaaagcgttgaagtcggaataaagtlatgaa 1800
   |||||
Db 1741 GAAATAACATTAATTAATCAAAATTAATAAAAGCGTTGAAGTCGATTAAGTATTTGA 1800
QY 1801 ttaagcactgttlaacgagaacacataatgattcaaaaactcaaaacttaactgtcttaca 1860
   |||||
Db 1801 TTAGGCACTGTTAACGAGAAACAAATATGAGTTCAAAACCTCAAGACTTAATCTGCTTACA 1860
QY 1861 gttattagcgtlaacgcttcaaaaatgycgacgctgtgaattaaagtaacgtgaactcaaga 1920
   |||||
Db 1861 GTTATTAGCGTAACTGTTAAATATGCGGACGCTGATTTAAAGTAACTGAAGCTCAAGAA 1920
QY 1921 ttactgtgaagttctcaagaatttaataacttaaatgtctacaacgctttcggttagc 1980
   |||||
Db 1921 TTACTGTGAAGTTCTCAGAGAAATTTAAATTAATTAATGCAACACCGTTGCGGTAGC 1980
QY 1981 acaatcacatagcgtlaaagttgtgtgaatgaagcgtgtgcaacttatctgtcttaca 2040
   |||||
Db 1981 ACAATCACATACGCTCAAGTGTGCTGTAGTAAAGCGGAGCAAACTTATCTGCTTTACA 2040
QY 1981 gtaagtgatcgtgttcaacaaatagaagcgttaacaaaggttacaatatagtagtgcgt 2100
   |||||
Db 1981 GTGAAAGTTGCTCTTACCAAAATTTAAGCAATTAACCAAGGATGATCAAAATTAAGTTCGCT 2100
QY 2101 aaagtgtcaacagctcctgtttaaagtgctgcgaactgtgcaaatctttagcaactaact 2220
   |||||
Db 2101 AAAGTGTCAACAGCTCTCTGTTAAAGTGTGCAAAATGCAAAATCTTATAGCAACTAATCT 2220
QY 2161 aaagtgtcaacagctcctgtttaaagtgctgcgaactgtgcaaatctttagcaactaact 2280
   |||||
Db 2161 AAAGTGTCAACAGCTCTCTGTTAAAGTGTGCAAAATGCAAAATCTTATAGCAACTAATCT 2280
QY 2221 atctatactaltaacactgaaggtcaagcgtlaacagcacaacaaaggttacaagaatgac 2280
   |||||
Db 2221 ATCTATACATTTTACACTGAAGGTCAAGGCTAAAGCAAGCAACGCTTACAAAAGATATTC 2280
QY 2281 aaagtgtatctttaaagaacgctgtgtagtcaacttaacaaactgaagcgttgaatgagct 2340
   |||||
Db 2281 AAAGGTGATTCCTTTAAAGACGCTGATGCAAGTTACTACACTTAACGAACCTTGTATGACGT 2340
QY 2341 caaaaattacatccaatttagcgaagaattaaaactttagtggtctttagtggct 2400
   |||||
```

```

Db 2341 CAAAAATTCATCCATTCATTCAGAGAGATTAAAACTTCTAGTGTCTTATGAGGT 2400
QY 2401 ggcacaaagtaactgctgagagaattaaacaaacgaatgggtgtagtgcgtctcgaaga 2460
Db 2401 ggcacaaagtaactgctgagagaattaaacaaacgaatgggtgtagtgcgtctcgaaga 2460
QY 2461 actgtatcagttgctcctcgaagacagatgcaatgggtaagtaacagctcgtgtgtaca 2520
Db 2461 ACTGTATCAGTTCGTCCTAGACAGATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2520
QY 2521 ttaactggtcctcgaagacagatgcaatgggtaagtaacagctcgtgtgtaca 2580
Db 2521 TTAACGTGCTCTGACAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2580
QY 2581 tctactgtaggaattgctgtagttagttagttagttagttagttagttagttagttagt 2640
Db 2581 TCTACTGATGAGAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
QY 2641 cgttacacagctgtagacacactgtagctcgtgcaaacgctgctgctgcaaacagatgt 2700
Db 2641 CGTTACACAGCTGAGACACACTGATGCTGCTGCAAAAGCTGCTGCTGCAAAAGCTGCT 2700
QY 2701 caaacgctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 2760
Db 2701 CAAAACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2760
QY 2761 gttgattcaatgaacacatgatttagcgaagttaaacctggaacacacgtctgtttaa 2820
Db 2761 GTTGAAATTCATGAAACATGATTTAGCGGAGTTAAACCTGGAACACATGCTGTTTAA 2820
QY 2821 ggcagcagcgaatgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtgtagcgtgtag 2880
Db 2821 GCAGCAGGTAATGCGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
QY 2881 gttactcactcactcactcactcactcactcactcactcactcactcactcactcactcact 2940
Db 2881 GTATTCTACTCCATCTCAAAATTTAAAGTGGTACAGTTACTCTGTTAACTGACGGT 2940
QY 2941 gtagagataagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtagtag 3000
Db 2941 GTGAGAGATTAAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 3000
QY 3001 ggcgaatcgaacgttatcctcactcactcactcactcactcactcactcactcactcact 3060
Db 3001 GCGAATCCAACTGATCTTCAATTCAGCATCTGCTGAGCGGTGAGTAACTGACCGTCT 3060
QY 3061 aaaaacatatacattgaattcagcagatcagcagatcagcagatcagcagatcagcagat 3120
Db 3061 AAAACAAATTCATGAAATTCAGCATCTGCTGAGCGGTGAGTAACTGACCGTCT 3120
QY 3121 gctgacggaactcacttactactactactactactactactactactactactactactact 3180
Db 3121 GCTGACGGAACTTCATTTACTAATTCACATTAGTAAATGTAATGTAATGTAATGTAAT 3180
QY 3181 tacaacaaatgtagtccacaaaggtgtagaactgtagaactgtagaactgtagaactgtaga 3240
Db 3181 TACAACAAATTTGATTCACAAAGGTGTAACACTGAGTGTACTTACTCAATATGATGATG 3240
QY 3241 gttcaaaaagattcctcaaacgtgtagttagttagttagttagttagttagttagttagt 3300
Db 3241 GTTCAAAAAGATTTTCAAACTGTACTGATATTGATGCAAAAGTTACATTCATCCAGGT 3300
QY 3301 tctgttgcctacgacgaagtaaacacgtcctcgtagtagcgtgtgtgtgtgtgtgtgtgt 3360
Db 3301 TCTGTGCTACTGACGAGAGTAACCTCTGCTAGAGGCTTGTGTGTGTGTGTGTGTGTGT 3360
QY 3361 agctatcactcagatgctgtagcgaacacgactcgtgtgtgtgtgtgtgtgtgtgtgtgt 3420
Db 3361 AGCTATACTCAGAGAGCTGACGACGACGACTGCTGTGTGTGTGTGTGTGTGTGTGTGT 3420
QY 3421 ccaagttgcctcctcaattcctcagaaggtatcgtatcagaagtaacgacgtgtagcagta 3480
Db 3421 CCAAGTTGCCCTCAATTTCTCAAGAGGTATGATTTAAAGCAATGCAATGCTGACAGTAACA 3480

```

```

QY 3481 aatattactgtagataaaactgttgaagttatttcaaaaagagagtgtagacgacacat 3540
Db 3481 AATATTACTGATGATTAATACTTTGAGATTATTTTAAAGAGAGAGTGTAGCCACACCAT 3540
QY 3541 gatcaggtgctcactcaagagacatgtagtaataacacagctacccctttagttagttagt 3600
Db 3541 GATGACAGTGTCTACTAGAGAGACATTAGTAATTAACACAGTGTAGCTTACTTACTTGA 3600
QY 3601 aacagcaagcttataagattgttgaagttgtagttagttagttagttagttagttagttag 3660
Db 3601 AACAGCAAGCTTATTAAGATTGTTGTAAGTGAATTAAGAGCAGCAGGTAAATGTTGCA 3660
QY 3661 gatattactatcattatataaagtaa 3687
Db 3661 GATTAATTACTATTCTATATTAAAGTAA 3687

RESULT 4
BSSBSA
LOCUS
DEFINITION B.stearothermophilus sbas gene for S-layer protein (sbas).
VERSION X71092
KEYWORDS cell envelope protein; cell membrane associated protein; S-layer protein; sbas gene.
SOURCE
ORGANISM Bacillus stearothermophilus.
REFERENCE
AUTHORS Kuen, B., Sleytr, U.B. and Lubitz, W.
TITLE Sequence analysis of the sbas gene encoding the 130-kDa surface-layer protein of Bacillus stearothermophilus strain PV72
JOURNAL Gene 145 (1), 115-120 (1994)
MEDLINE 94320770
REFERENCE 2 (bases 1 to 3832)
AUTHORS Kuen, B.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-1993) Kuen, B., Institute of Microbiology and Genetics, University of Vienna, Dr. Bohig, 9, 1030 Vienna, Austria
COMMENT
Related sequences: D90050 and Gilmoe, R.D., Mol. Microbiol., 6:1579-1586 (1992).
FEATURES
source
1..3832
/organism="Bacillus stearothermophilus"
/strain="PV72"
/db_xref="taxon:1422"
prim_transcript <1..3767
gene
1..3767
/feature="sbas"
CDS
37..3723
/gene="sbas"
/codon_start=1
/translation="MDRKKAYKLTPASGIAASAVYANPNANSEATDVAVYSQAKQ
FKKAYVYVSHVTETGEPTINDVYAEENAKRRYRAVALVNAAGAKDAYADADQ
KEYETVYFKAMPKSGEARVAVYIDAVYATFLIDEMROELEAVOAKDLKAEQYHKI
PYEIKTRVILIDRYGKTRDLRSTFKAKQELRDSLIYDITVAMKAREVDAAKRG
NLDKAKAAVDDINOYLKRVDAEFTELFEVAKKRLDADDEALTPKVSVAINTONKA
VELTAVPNGLIKIQLSNAAMEDVYVNTVYIYVDGNIIPALMTAVYSISTDCKTT
VDASTPEENTEYKVVVKGIDKNGKEKEDEAFKKLRNDVAVTQVEGTWNTSVN
LAAGTPEIDDTLVVDPKLABETVNSNVTITDVEGKRIPVIASTSGSTITLLE
ALVYGKOYKLAINNVTLTGYNAAVELYFANASAPVATAPTLGSTLSTSLTT
NMWGLAGVNEAGTYVPGLOFTTTPATKIDESTLADNPVAVKESGVVYASLKYNA
DAKWTIYVPAKADIENTIYQIKIKGLSKDGIELGVNKTVEFKRODILAPVIVIS
TSKNGDLKLTVEDEFTVKSSENLNFTNATVSGSTTTTGQVAVVAGAWLSLIVS
DIIPASVEAVTGDGTYKRVKVAANLERNOGYKLVEFGKATAPVKDAVANNTLATNV

```


YVTTTBEQDVTAPVTVKFKDLSLKDAVATLTLNVDAQKFTIQFSEELKTSQSL
VGVKATVBEKLTNGWVDAQTGVTSVAPKTDANGVTAIVVLITGDNNDKAKRLV
VDRKSTDEIADVAGNVIKERDILIRNSRHTVSAKAAADOGSAAPFSTAD
TKSLVEEDNDLAEVDEKPEVNVDAAGNAVATVTLDSGTNKFVFPSPSELKAGT
VYSVTLIDGVNRDQVGVTSIKYTSFTKSVSNPTLSSISLADGAVNDRSKTTLIESDS
VNPPTITLKKADGTSFTNTLVNHNENKTKIVHKGTLDLEFQYELAVSKDFQIG
TDIDSKVTFITGSAVTDKVPALVGVSGSNSTYQDAATATRSVADVAPVPLQF
SEGIDLTNAVTNTTDDKTEVVISKESVSDADHDAGATKETLVINTVTPVLVDSNKT
KIIVSGVKDAAGNVADITTFYIK"
37. 126
sig_peptide /gene="sbsa"
mat_peptide 127. 3720
/gene="sbsa"
/product="surface layer protein"
terminator 3767. 3806
BASE COUNT 1355 a 686 c 749 g 1042 t
ORIGIN

Query Match 100.0%; Score 3687; DB 2; Length 3832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atygataggaaaaaagctgltgaactagcaacagcaagtcgtatgcaagtcgatt 60
DB 37 ATGAGTAGGAAAAAGCTGTGAANAATAGCAACAGCAAGTGTCTATGCGACGACGATT 96
QY 61 gtcgctgcgaatcccaagccttcgaagcgcgtacagatgtagcaacagtagtaagccaa 120
DB 97 GTCGCTGCGAATCCCAAGCCCTTCTGAAGCGGCTACAGATGTACCAACAGTAGTACCCAA 156
QY 121 gcaaaagcagcttcaaaaagacatactactacagcacaatacagtaacggaactgct 180
DB 157 GCMAAAGCAGCTTCMAAAAAGCATCTACTTACAGCCATACATAACGGAAATGGGT 216
QY 181 gaattcccaacatlaagatgatatgctgaatacaacaagaacgaaaaacgatacgt 240
DB 217 GAATTCCCAACATPAAACGATGTATGTCTGAATACCAACAAAGCAAAAACGATACCGT 276
QY 241 gatgagtgatgacttagtgaataaagcaggtgcgcgaaaaaagaagcgttaactgactg 300
DB 277 GATGGCGGTAGCATAGTGAATAAAGCAGGTGGCGCAAAAAGACCTTACTTAGCGTGAT 336
QY 301 ttacaaaaagaatagaactacgtttccaaagcaaacccctaactctggcgaagctcgt 360
DB 337 TTACAAAAAGATATGCAAACTTACGTTTTCAAAAGCAAAACCTTAATCTGGCGAAGCTCGT 396
QY 361 gtagcaacttacatcgatgcttcaacaacta tgaacaaaaa tgaacgaaatgcccagaag 420
DB 397 GTAGCAACTTACATCGATGCTTACACACTATGCAACAAATTTAGACAAATGGCCCAAGAG 456
QY 421 ctaagagcgtgcgttccaagaacaaagattagaanaagcagaacaatactatacaaaatt 480
DB 457 CTAGAGGGCTGCTTCAACCAAAAGATTGAAAAAGCAGAACCAATCATCTCAAAATTT 516
QY 481 cctttagaattaaactcgacagtcatttagatcggtatagtggttaaaaactcgt 540
DB 517 CTTTATGAATAATTAANAACCTGCACAGTCATTTTAGATCGGCTATATGTAANAACCTCGT 576
QY 541 gatttactcgtctacattttaaagcaaaagcacaagaacttcgcgcagcgttaattat 600
DB 577 GATTACTCTCGCTTACATTTTAAGCAAAAGCACAAGAACTTCGCGACAGCTTAATTAT 636
QY 601 gatattaccgttgcgaatgaagaagcgcggaagtacagaagcgtgtgaagcaggcaattta 660
DB 637 GATTATTACCGTTGCAATGAAGCGCGGAAGTACCAAGACGCTGTGAACACAGCAATTTTA 696
QY 661 gacaaagctaaagcgtcgttgaatcaaatcaatcaactacttcaacaaagtaacagatgct 720
DB 697 GACAAAGCTAAAGCTGCTTGTATCAATCAATCAATTACTTACCAAAAAGTAACAGATGCT 756
QY 721 ttcaaaactgaactaacagaagtagcgaaaaaagcattagatgcagatgaagcgtgcctt 780

DB 757 TTCMAAACTGMACTAACAGAAAGTAGCGMAAAAAGCACTTAGANTCGAGATGAAGCTCGCTT 816
QY 781 accccaaaagtgtgaaggtgaagtgcgatttaacacactcaaaacaaagcgtgtgaattaca 840
DB 817 ACTCCAAAAGTTAAAGTGAAGTGCATTTAACACTCAAAACAAAGCTGTTGAATTTACA 876
QY 841 gcaatccagtgaaacggaacataaaatlacaacttcagctgcgtgcaaa tgaagatata 900
DB 877 GCAGTACCACTGAACGGAACACTAAATTTCAACCTTTACACTGCTGCAAAATGAAGATACA 936
QY 901 gtaacgtaaaatactgtaagtcataataaagtggaaggtataacatltgcccttaat 960
DB 937 GTAAAGCTAAATACGTACTATCTATTAAGGTGACGGTAACTTCATTTGCCCTTAAT 996
QY 961 acggcagatgttcttattatcacagcggaaaacatacactctgtgaatgttcaactcca 1020
DB 997 ACGGCAAGATGTTCTTATCTTACAGACGGAATAACTATACCTGTGGTGCCTCAACTCCA 1056
QY 1021 ttccgaataatacaggagata taaagtagtagttaaagata ttaaagcaaaaatgycgaa 1080
DB 1057 TTGCAAAAATATACGGAGTATTAAGTAGTATTAAGGTATTTAAAGCAAAAATGGCAAA 1116
QY 1081 gaattaaagaagatgcattcacttccaagcttcgaaatgaatgacgtgtagtlaactcaagtg 1140
DB 1117 GAATTTAAAGAAGATGCATTCACCTTCAAGCTTCGAAATGATGCTGTGATCTCAAGTG 1176
QY 1141 ttggaacttaatgtaacaaacaaactctcgttaacttagcagcaggtlaacttcgacact 1200
DB 1177 TTTGGAACTAATGTATTAACAAACAACTTCTGTAAACTTACACAGGTACTTTCGAACT 1236
QY 1201 gacgatactttaaagtagtatttgaatgaatgtttagcaccgttgaactgttaaacagctcg 1260
DB 1237 GACGATACCTTAAACAGTAGATTTGATTAAGTTTGAACCTTAAACTGTAAACAGCTCG 1296
QY 1261 aacgttactttaaacagatgttgaactggaacacgaatccagtaattgtactactctc 1320
DB 1297 AACGTTACTTATACAGATGTGTAACCTGGAACCGCATTCGAGTAATTTGCATGCTACTTCT 1356
QY 1321 ggttctacaacttaacttaacgtttaaagaagcgtttagtgaactgtgtaacaaatataaact 1380
DB 1357 GGTTCACCAATTAATCTATTAAGTTAAAGACGCTTGTGTAACCTGTGTAACCAATTAACCTT 1416
QY 1381 gctacaataatgttlaaaaacatlaactggttacaatgacagaagcgttlaacgattggttc 1440
DB 1417 GCTATCAATATATTTAAACATTAACGTTACAAATGCAACACTTACGAGTTAGTTC 1476
QY 1441 actgcaaaacgcacacagcaacactgttgcaccgctcctaacttacttaagtggtlacaact 1500
DB 1477 ACTGCAAAAGCATACAGCAACACTGTTGCTACCGCTCTACTACTTATAGGTGTACAACT 1536
QY 1501 ttacttactggttctcttacaacaaatggttgggttaaatgttgctggtgtgtaatgaa 1560
DB 1537 TTATCTACTGCTTCTCTTACAAACAAATGTTGGGGTAAATTTGCTGTGCTGTGAATGAA 1596
QY 1561 gctggaacttatatcctcgttcttcaatlcacaacaacgcttgcctaaggttlaagcgaa 1620
DB 1597 GCTGGAACTTATTAATCTGCTTCAATTCAACAACAGCTTCTACTAAGTTAGACGAA 1656
QY 1621 tctacttagctgataactttgatttagtgaaaaaaagatcgtgatacgttgttgcctc 1680
DB 1657 TCTACTTTAGCTGATTAATCTTGTATTAGTTGAAAAAAGAACTCGTACAGTTGTTGCTTCT 1716
QY 1681 gaactaaatataatgcagagcgttlaaaatggttaactttagtgcctaaagcggaccttaa 1740
DB 1717 GAACCTAAATATTAATGACAGCGCTAAATGTGTAACTTTAGTGCCTCAAAAGCGGACCTTAA 1776
QY 1741 gaaaatacaatctatcaaatcaaatctaaaaaagcgttgaagcgcgataaagtatgtga 1800
DB 1777 GAAATATCAATCTATCAATCAAAATTTAAAAAAGCGTTGAAGTCGATTAAGTATTTGAA 1836
QY 1801 ttaggcactgtttaaagagaacaataatgatttcaaaaactaaacttaactgtccttaca 1860
DB 1837 TTAGGCACCTGTTAAACGAGAAAACATATGAGTTCAAAAACCAAGACTTAACCTGCTCTACA 1896

Oy	1861	gtatttagcgtacgtctcaaaaatgacgccttgattaaagaactgaagctcaaga	1920
Db	1897	ggtatttagcgtacgtctcaaaaatgacgccttgattaaagaactgaagctcaaga	1956
Oy	1921	ttactgtgaagctctcagaagaattaaatacatttaactcacaacgcttcgggtagc	1980
Db	1957	tttactgtgaagctctcagaagaattaaatacatttaactcacaacgcttcgggtagc	2016
Oy	1981	acaacacatacagctcgaagcttgctgtagtaaaagcgggtgcaaaccttctcttaca	2040
Db	2017	acaacacatacagctcgaagcttgctgtagtaaaagcgggtgcaaaccttctcttaca	2076
Oy	2041	gcaagctgacatccatccagctagctgttgaagcgttactgctcagaagatgaaacata	2100
Db	2077	gcaagctgacatccatccagctagctgttgaagcgttactgctcagaagatgaaacata	2136
Oy	2101	gtgaaagctgctcgtcaacccaattagaacgttaaccaaagggtacanaattagtagtgcgt	2160
Db	2137	gtgaaagctgctcgtcgaacccaattagaacgttaaccaaagggtacanaattagtagtgcgt	2196
Oy	2161	aaagctgcaaacgctccctgttaaaagatgctgcaaatgcaaaactcttgcaactactat	2220
Db	2197	aaagctgcaaacgctccctgttaaaagatgctgcaaaactcttgcaactactactat	2256
Oy	2221	atctaatacttacaactgaaggtcaagaacgttaacagcaaccaacggttacaagaatctc	2280
Db	2257	atctaatacttacaactgaaggtcaagaacgttaacagcaaccaacggttacaagaatctc	2316
Oy	2281	aaagctgtagctcttcaaaagacgtgtagtactcaactcaagcaagctgtagtgcagct	2340
Db	2317	aaagctgtagctcttcaaaagacgtgtagtactcaactcaagcaagctgtagtgcagct	2376
Oy	2341	caaaatcactatccaaatlltagcgaagaataaaactcagctggtctcttagtggt	2400
Db	2377	caaaatcactatccaaatlltagcgaagaataaaactcagctggtctcttagtggt	2436
Oy	2401	ggcaaaagtaactgtcgaagaatlaaacaacgaacgtgtggtagctgtgtagtgcgaaca	2460
Db	2437	ggcaaaagtaactgtcgaagaatlaaacaacgaacgtgtggtagctgtgtagtgcgaaca	2496
Oy	2461	actgtagaagctgctccctaaagacagatgcaaatgtgtaagtaacgctgctgtagtaca	2520
Db	2497	actgtagaagctgctccctaaagacagatgcaaatgtgtaagtaacgctgctgtagtaca	2556
Oy	2521	tttaactggtctctgacaataacacgaacgaatgcaaatgctgctgtagtacaatgct	2580
Db	2557	tttaactggtctctgacaataacacgaacgaatgcaaatgctgctgtagtacaatgct	2616
Oy	2581	tctactgtagaattgtcgtgtagtgcgtgtaattgaatgaagaaagaatactttaat	2640
Db	2617	tctactgtagaattgtcgtgtagtgcgtgtaattgaatgaagaaagaatactttaat	2676
Oy	2641	cgttacacacagctgtagacacacacgttagcttctgtgaaagcgcgcgtgcgaacaagatgct	2700
Db	2677	cgttacacacagctgtagacacacacgttagcttctgtgaaagcgcgcgtgcgaacaagatgct	2736
Oy	2701	caaaacgctcgtcgtgcaatcccaacaacgactgcaattgatacactaagaagcttata	2760
Db	2737	caaaacgctcgtcgtgcaatcccaacaacgactgcaattgatacactaagaagcttata	2796
Oy	2761	gttgaatcaatgaactgatttagcgggaagtlaaacctggaagacatcgtgttlaaagat	2820
Db	2797	gttgaatcaatgaactgatttagcgggaagtlaaacctggaagacatcgtgttlaaagat	2856
Oy	2821	gcacagagtaatgctgtagctgctgtagtcttaacagcactaagacggttctcaataatct	2880
Db	2857	gcacagagtaatgctgtagctgctgtagtcttaacagcactaagacggttctcaataatct	2916
Oy	2881	gtatcactcactcactcaagaattaaagcgtgtacagttactcgttaacaaattagct	2940
Db	2917	gtatcactcactcactcactcaagaattaaagcgtgtacagttactcgttaacaaattagct	2976

Oy	2941	gtgagagataaagtagtgaacacaaatctcctaaatacattactcgttcaagaatctatc	3000
Db	2977	gtgagagataaagtagtgaacacaaatctcctaaatacattactcgttcaagaatctatc	3036
Oy	3001	gcgaatccaaagcttatccttcaatcagcatctgtagcgggtgcaagcttgaacgcttct	3060
Db	3037	gcgaatccaaagcttatccttcaatcagcatctgtagcgggtgcaagcttgaacgcttct	3096
Oy	3061	aaaacaattacatggaattcagcgatctcagttccaaaccccaacaatacctttaaga	3120
Db	3097	aaaacaattacatggaattcagcgatctcagttccaaaccccaacaatacctttaaga	3156
Oy	3121	gctgacggaactcacttaactaactagtaaatgtaaatatgaataataaaca	3180
Db	3157	gctgacggaactcacttaactaactagtaaatgtaaatatgaataataaaca	3216
Oy	3181	tacaaatgtatcccaacaaggtgtaaaccttgacgagttactcaataaagattaga	3240
Db	3217	tacaaatgtatcccaacaaggtgtaaaccttgacgagttactcaataaagattaga	3276
Oy	3241	gtttcaaaagatttcaaacctgtagtgaatgtagaagaattgaatcaacagct	3300
Db	3277	gtttcaaaagatttcaaacctgtagtgaatgtagaagaattgaatcaacagct	3336
Oy	3301	tctgttctactgacgaagtaaaacctcgtcctagtaggacgtgtgttcaatgaaac	3360
Db	3337	tctgttctactgacgaagtaaaacctcgtcctagtaggacgtgtgttcaatgaaac	3396
Oy	3361	agctatctcagagatgctgcgacgaacacgacttggtctgttagcgtgacttgctcgag	3420
Db	3397	agctatctcagagatgctgcgacgaacacgacttggtctgttagcgtgacttgctcgag	3456
Oy	3421	ccagctgcccctcaattcctcaagaaggtatcgatlaacgaagtaacactgtagaataca	3480
Db	3457	ccagctgcccctcaattcctcaagaaggtatcgatlaacgaagtaacactgtagaataca	3516
Oy	3481	aattactgataataaactgttgaagttatctcaaaagagaatgtagaagcagcaat	3540
Db	3517	aattactgataataaactgttgaagttatctcaaaagagaatgtagaagcagcaat	3576
Oy	3541	gatgagagtgctactaagaagacatagtaattaaacacagctaccccttagttagt	3600
Db	3577	gatgagagtgctactaagaagacatagtaattaaacacagctaccccttagttagt	3636
Oy	3601	aacaagaagacttaataaagttgtgtaagtgtagtlaaagaatgcaagcaggtlaattgca	3660
Db	3637	aacaagaagacttaataaagttgtgtaagtgtagtlaaagaatgcaagcaggtlaattgca	3696
Oy	3661	gatactatacatctataataagtaa	3687
Db	3697	gatactatacatctataataagtaa	3723

RESULT 5

AF055578	3680 bp	DNA	BCI	08-MAR-2000
LOCUS				
DEFINITION	Bacillus stearothermophilus surface layer protein precursor (sbpc) gene, complete cds.			
ACCESSION	AF055578			
VERSION	AF055578.2	GI:4581955		
KEYWORDS				
SOURCE	Bacillus stearothermophilus.			
ORGANISM	Bacillus stearothermophilus			
	Bacteria: Firmicutes: Bacillus/Clostridium group; Bacillus.			
REFERENCE	Bacillus/Staphylococcus group; Bacillus.			
AUTHORS	I (bases 1 to 3680)			
TITLE	Jarosch, M., Bgelsee, E.M., Maltanovich, D., Sleytr, U.B. and Sara, M.			
	S-layer gene sbpc of Bacillus stearothermophilus ATCC 12980:			
	molecular characterization and heterologous expression in			
	Escherichia coli			
JOURNAL	Microbiology 146 (Pt 2), 273-281 (2000)			
MEDLINE	20170659			
PUBMED	10708365			

REFERENCE 2 (bases 273 to 3572)
AUTHORS Jarosch,M., Egeleer,E.M., Mattanovich,D., Sleytr,U.B. and Sara,M.
TITLE Direct Submission
JOURNAL Submitted (25-MAR-1998) Zentrum fuer Ultrastrukturforschung,
Universitaet fuer Bodenkultur, Gregor Mendelstrasse 33, A-1180
Vienna, Austria
3 (bases 1 to 3680)
AUTHORS Jarosch,M., Egeleer,E.M., Mattanovich,D., Sleytr,U.B. and Sara,M.
TITLE Direct Submission
JOURNAL Submitted (13-APR-1999) Zentrum fuer Ultrastrukturforschung,
Universitaet fuer Bodenkultur, Gregor Mendelstrasse 33, A-1180
Vienna, Austria
REMARK Sequence update by submitter
COMMENT On Apr 13, 1999, this sequence version replaced gi:3025825.
FEATURES
source Location/Qualifiers
1..3680
/organism="Bacillus stearothermophilus"
/strain="ATCC12980"
/db_xref="taxon:1422"
/db_xref="ATCC:12980"
51..3643
/gene="sbsc"
51..3643
/gene="sbsc"
84
/gene="sbsc"
/note="alternative transcription start site"
255..261
/gene="sbsc"
273..362
/gene="sbsc"
273..3572
/gene="sbsc"
/note="S-layer protein: Sbsc"
/codon_start=1
/transl_table=11
/product="surface layer protein precursor"
/protein_id="AAC12757.1"
/db_xref="GI:3025825"
/translation="MDKKKVKLATAASAVASAFPAANPHASQAATDVATVVSQKAO
MKEATYTSHTVETGQPDINKDVTAATNKAKQATVANAIVAVNKAAGKADYIADLIQ
AIETYYVRKAPKSGEASAVAYIDAVYATATKDKROELKAADVADKADKAEELYHI
SYELKTRVILIDRVYGSTRELIRSTFAKADQALDRILYDITVPMKAREADADKAG
NLDRKAKLDVNOYVSVKVFDAFKAELQKADODAAEAALTPREVESAIQDSNFK
VTFKPVDKATAIPKNSITLTKGPKTKPKSPKVSFSGSGITPVTIYPTINDKRYTV
VTSGLKDTAGKEFEISTNEFTYNNKVPASITFNNKLEPDSAVDITKTYTVKDAAGNV
IKSGELEFTSSEKLTGKFLNTGKKSIVYNAVTVKGNVTGNVILAVEDEKAASVS
ELKLRDKRKEVATILYANGNAPDKDGNQISGTLTLTAKFKDQYGNELGKAGDTYTE
ESLNFEVLVAVPDGSVTPIDPGTALVKYKGEVETIIPVTVKANPVLNETIADVSTGS
VAKGOKAFPKYTLKDOYGNKFTGNVNTSDKETETAVSNGSIGOSEFTVYVNGVAG
GSTTITKSGTRKVVYPVNVVAGGPVAVQIKVLDGKIDKASATSPANNVQLVYA
VDANGNIYGDITNDYITTSSEATDITNGVYVNAKSTANGDTYVYITDNSKKRYKETLL
VKLGTVILGTVDEVIDTTLKATVATKADLLEDAUNDGDLAKLANLADIKDNGN
PMVBSAATPNTNEKLQALKSVLSGIVSSDVTSGSVNDNLKDAISGLAVKAKGT
VTLVLFNEDSKRIAPAITVKAIPATQGVTVTGDLDPGVYGVKTFETADIKSG
HKLYAVDSDAVPAPAVGTTRNSTFEANIEVGTGEVAVANAOIITVIEVSDSVK
KTFTEVAADLSVADKRTGPTATVPTGNGOYTKETLLAVNDGLANGKIXVAAAGSS
AAAPVKGATYDITVTRTYYTGGTGVTSIGVPEVDAQGHIISIEVDENKGVYKDYTI
TGIOIGIKRSAS"
363..3569
/gene="sbsc"
/product="surface layer protein"
3606..3643
/gene="sbsc"
BASE COUNT 1265 a 694 c 766 g 955 t
ORIGIN
Query Match 13.9% Score 513; DB 1; Length 3680;
Best Local Similarity 66.6%; Pred.No.7.7e-79;
Matches 767; Conservative 0; Mismatches 375; Indels 9; Gaps 2;
Qy 1 atgataagaaaaagctgtgaactagacacagaactgtctatgcagcaagtgcatctt 60

Db 273 ATGCACAAAAAGAAAGCAGTCAACTCCGACAGCCAGCCCTTAGCAGGACGCCCTTTC 332
Qy 61 gtgcgtgcaaatccaaacgcttcgtgaagcggctacagatgtagcaacagtagtaagccaa 120
Db 333 GTCGACAGCCAAACCCGATGCTTCCCAAGCAGCAACGAGCAGTGGCGATGTCACAGCA 392
Qy 121 gcaaaagcagatgctcaaaaaagcattactactactactacagcattacaggaactggt 180
Db 393 GCGAAAAGCCGAATGAAAGAGCATCTATACCTACAGCCATCGGTAAAGGAAAGCGGC 452
Qy 181 gaattcccaaacattacagatgctatgctgtaatacaacaagcggaaaaaagataccgt 240
Db 453 CAATTCGCCGACCTTAAAGATGTATATCCGCCCTACACAAACAAACAAAGCGTTATGCA 512
Qy 241 gatcgtagtgcattagtgaaagcaggtgycgcgaaaaaagacgcttacttagtcat 300
Db 513 AACGCGGTTGCTGTCTGAAACAAAGCGGGGCGCGCAAAAAGACGGATTTTGCTGAT 572
Qy 301 ttcaaaaaagaatagaactacgcttccaagcacaacccctaactctgcygaactgct 360
Db 573 TTGCAAGCTATATATACGAAAGCATGTATTCAAAGCCAACCCGAAATCGGCGAAAGCGCG 632
Qy 361 gtagcaacttacatcgatgcttacaactatgcaacaataattagacgaatgcgcaag 420
Db 633 GTGGCTACGTCATCATCGACGCTTCAACTATGCGACAAATATACACAAATGCCCAAGAA 692
Qy 421 ctagaagctgctgttcaagcaaaagattagaaaaagcagaaacatactatcaaaaatt 480
Db 693 CTCMAAGCAGCCGTCGATGCGAAGATCTGCAAAAAGAGAGAGCTCACCAAAAT 752
Qy 481 ccttagaataatgaactgcagacagtcattttagatcgcgtatagtgttaaaactcgt 540
Db 753 TCGTACGAGCTCAAAAGCGCAGCGATCTTCGACCGCTACGCGCAATGACGCGC 812
Qy 541 gattactcgcctctacatttaagcaaaaagcacaagaacttcgcagacgcttaattat 600
Db 813 GAACCTCTCGCTTCGACGCTTCAAAAGCGCATGCTCAAGCATGTGGCGACAGATTGATCTAC 872
Qy 601 gatattaccgtgtgcaatgaagaagcgcggaagtagaagaagcgtgtgaagagcaatt 660
Db 873 GACATACCGTTTCCATGTGAAGCGCGTGAAAGCCCAAGACCGCTTCACAGCCGCAACTTG 932
Qy 661 gacaaagctaaagctgctgtgtgaatcaatcaactacttaccaaaagtaagaagtgtc 720
Db 933 GACAAAGCGAAACCCCGCTTGACCAAGTCACCAATAGCTATGCAAGTAACCGATGCG 992
Qy 721 ttcaaaactgaactaacagaagtagcgaaaaaagcattagatgcagatgaagctgcct 780
Db 993 TTCAAAGCGCAACTGCAAAAAGCAGCAACAAGACCGCAGCGCATGACAGCAGCAT 1052
Qy 781 aactcaaaagctgaagaagctgaagctgcattacaactcaaaaacaagcgttgaatta 840
Db 1053 ACCCCCAAGTTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1112
Qy 841 gcaagtagcagtgtagcgaacacataaataacttaacatttcagctgcgtgaatgaag 900
Db 1113 AGCCCTGTGACCAACGACCAAGCAATTCCTTAAACCTTCTTATTAAGTTAAAGGATCT 1172
Qy 901 gtaacgtaatactgttagctatctataaagtggaagcgttaaacatttccttaatt 960
Db 1173 GAAA---CTAAACTTATATCTTAAGCTGTGTTGAAGTATCGAGAGTGGATTAACTGCTACT 1229
Qy 961 acggcagatgcttcttactactacagacggaaaaaactacaactgtagtgcctcaactc 1020
Db 1230 GTAACTCTTATGATACATTAGTAGACGGTAAACCTTACACTGTTGAACATCTGCT--- 1286
Qy 1021 ttgcgaataataatagagagataaagtagtagttaagagataaagataaagataaagtgca 1080
Db 1287 ---TTAAAGATACACGCTGTAAGATTTGAAGATTGAAATAGCAACAAGCAATTCATCAAC 1343
Qy 1081 gaattaaagaagatgcatcacttcaagcttcgaagcgaatgagctgttagtactcaagct 1140

Db	1344	AMGCAGTACACGCTTCATATGCACTGCATTAAGTTACCGAGATGTCAGTT	1403
Oy	1141	tttgaactaa	1151
Db	1404	GATTGACTAA	1414
RESULT	6		
LOCUS	AF228338	3090 bp	DNA
DEFINITION	Bacillus stearothermophilus Sbsd precursor (sbsd) gene, complete cds.	BCP	24-JUL-2000
ACCESSION	AF228338		
VERSION	AF228338.2	GI:9186638	
KEYWORDS			
SOURCE			
ORGANISM	Bacillus stearothermophilus. Bacillus stearothermophilus Bacteria: Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.		
REFERENCE	1 (bases 1 to 3090) Danhorn,T., Egelseer,E.M., Idtis,R., Sleytr,U.B. and Sara,M. Molecular characterization of the first bacterial S-layer glycoprotein Sbsd from a temperature derived mutant of Bacillus stearothermophilus ATCC 12980		
TITLE	unpublished		
JOURNAL	2 (bases 1 to 3090) Danhorn,T., Egelseer,E.M., Idtis,R., Sleytr,U.B. and Sara,M. Direct Submission Submitted (24-JAN-2000) Zentrum fuer Ultrastrukturforschung, Universitaet fuer Bodenkultur, Gregor Mendel-Strasse, Vienna A-1190, Austria		
REFERENCE	3 (bases 1 to 3090) Danhorn,T., Egelseer,E.M., Idtis,R., Sleytr,U.B. and Sara,M. Direct Submission Submitted (14-JUL-2000) Zentrum fuer Ultrastrukturforschung, Universitaet fuer Bodenkultur, Gregor Mendel-Strasse, Vienna A-1190, Austria		
AUTHORS	Sequence update by submitter On Jul 14, 2000 this sequence version replaced gi:6984125.		
TITLE	Location/Qualifiers		
JOURNAL	1..3090		
REMARK			
COMMENT	/organism="Bacillus stearothermophilus"		
FEATURES	/db_xref="taxon:1422"		
source	255..261		
RBS	273..2984		
gene	/gene="sbsd"		
CDS	273..2984		
	/gene="sbsd"		
	/note="surface layer glycoprotein"		
	/codon_start=1		
	/transl_table=1		
	/product="Sbsd precursor"		
	/protein_id="AAF34763.2"		
	/db_xref="GI:9186639"		
	/translation="MDKKKAKKLTAASNAVASAAYAAPNPASQATDVATVYSSAKQAQ MKEAVYVSHVETGKPLNSDVAAYNAKQAYANAANAAGAKKADAYADIAQ ATYATVYKANKPKSGEARVATYIDAVYATGLDMRELKAPKADLKAELIYHKI SYELKTRTVILDRVYGOSTRELIRSTFKADQALRDSIYDITVAMKAREQAKAKG NLDKAKALDOVNOYVSRYMDPAFEALQKADQANAAVEALPKVSVYNAKKTLE IKPKKADVATVDNKTSPDSDVVKATITTLAIDQGSVSYVAKSLSDKTKLKL VADGSPFKRYVDIKNKRLDKDKVPSYTTTIDTDSVAPSVLSSVADNGLLTV KESPLASVGVKLYDGTTELISVSPKFTAGDEMTINLASSVYNDLITKIGADV YNGVAINPNPAELTVKRTVDYTPVQVINEAVNTKTVYFSEKLSNPYIKIGGOT ASVSDTGLVYATLANALSEGVAYEVSRYKDLACNAGDTYKVVOLKADTAPKRE VSSQVYKIDGVEHYVLTPEDEVYTGSNITVYQASDKYVDENNYLKVGLTTTSDNF KLYLPTGKSKSVANLNTSLPKGYTYTLPEDLVSDIAGNYAEKKOITVYRGSDSLT IKPLADSSYDNGVKADNNELVFAFQNDASLNSFNINGLAVTKAVFDSDIYH IRVPLAGANTWGTGTHVITISIKNTGLWMDYTVSEFKENAAPTFTALISADYV LDGSESEVANAATISSPLANNFTVQAGVYVAVYEDNNATVOCGSGCYKTVYLLK LQSPHSLSKPITLSAIGYIVDQGTGATNNNNLVGNVSNAAVYVAK"		
	273..362		
	/gene="sbsd"		
	3010..3050		
	terminator		

BASE COUNT	941 a	716 c	746 g	687 t	ORIGIN
Query Match	13.2%	Score 487.6;	DB 2;	Length 3090;	
Best Local Similarity	73.9%	Pred. No. 1.8e-74;			
Matches 619;	Conservative	0;	Mismatches 219;	Indels 0;	Gaps 0
QY	1	atggaatgagaaanaagctgtgaacttagcaacgaagctgtctatgtcagcaagtcattt	60		
DB	273	ATGGACAAAGAAAGCACTCAAACTCGCCACAGCCAGCGCTGTACGACGAGCGCTTTC	332		
QY	61	gtcgcgtgcaaatcccaaacgctctggaagcggtctacagatgttagcaacagtagaaaccaa	120		
DB	333	GTTCGGCGCCCAACCCCGCACCTTCCCAAGAGCAACCGAGCTGGCGCGGTCTGTCACCA	392		
QY	121	gtcaaaagcagcttcaaaaagcatcactactactaacgacatcagtaacgaaactgtgt	180		
DB	393	GGGAAAGGCCAATGAAAGACATCTCTATTACGTACAGCCACACGGTAAACGGAAACGGCG	452		
QY	181	gaattcccaacattacagatgtlatalgtctgtaatacaacaagaacgaaaaaacgatacgt	240		
DB	453	AAATCTCCCAACATCATGATGACGTATACGCCGCTTACAACAGCGCAACACGCGTATGGC	512		
QY	241	gattcggtgacatgattgtaataaagcaggtgtgcgcgaaaaaagaagccttactagtat	300		
DB	513	AACGGCGTCTGCTGTGTGTAACAAAGCGCGCGGCGCCAAAAGACCGCATTTGGCTGTAT	572		
QY	301	ttacaaaaagaatatgaanaactcagcttttcaagaacaaacctbaactcttggaagctgt	360		
DB	573	TTGGCAAGCTACATTACCAACCGTATGTATTCAAGGCCAACCCGCAATTCGGCGAAGCGCCG	632		
QY	361	gtgaacacttaactcgaatgtcttacaacta tgcgaacaaataggacgaaatgtgcgcaag	420		
DB	633	GTGCGACGCTACATTGTAGCCTTACAACCTATGCGACAAATTTAGCGCGATGGGTGAGAA	692		
QY	421	ctagagcgctgtcttcaagaacaaagatttagaaaaagcgaacaaatactatcacaatt	480		
DB	693	CTCAAAAGCGCCGTGGAACCGAAAGATCTGAAAAGAGCAAGAGCTCTACCAAAATTT	752		
QY	481	cctttgaataataaactcgcagatcatttagatcggtatagtgtaaacaactcgt	540		
DB	753	TGGTATGAGCTAAAACGGGACGGGTCATTTCTGACCGGCTCTACGGCCAATGAGCGCG	812		
QY	541	gattactcgcgtctacatttaagaacaaagacaagaactcgcgacagcttaattat	600		
DB	813	GAACTCTCTCCGTTTCGACGCTTCAAAAGCGCATGTGCTCAAGCATTTGGCGCACAGCTTGATCTAC	872		
QY	601	gatatcaccgttgcgaatgaagaacgcgcgaagtaacagacgcgtgtgaagcaggaattta	660		
DB	873	GACATCTACGTGTTGCCATGAAACGGCGTGAAGGCGCAAGAGCGCGCCAAACCGCAACTCG	932		
QY	661	gacaaagctaaagcgtcgtgttgaataaactcaatctcaacttaccaaaagtaacagtgct	720		
DB	933	GACAAAGGGAAGCGCGCTTGACCAAGTCAACCATATATCTTCGAAATTAACCGGATGGC	992		
QY	721	ttcaaaactgaactacaagaatgcgaanaaaagcatatgattgcagatgaagctgcct	780		
DB	993	TTCAAAAGCGCACTGCCAAAAGCAGCGCAAGACGCCAACGCGCGGTATGACACGCGGTTGG	1052		
QY	781	actccaagaagtgtgaagtgtaagtcgattacacttcaaaccaaaagcgtgttgaattaa	838		
DB	1053	CCTCCAAAAGTGAAGTGTATACAGCGCGTTACAGGCCAAACGTTGGAAATTCAAATTC	1110		
RESULT	7				
ID	E10125				
AC	E10125	standard;	RNA;	UNC;	2069 BP.
XX	E10125;				
XX	E10125.1				
XX					

Query Match	2.1%	Pred. 77.4%	DB 45%	Length 2069
Best Local Similarity	43.0%	Pred. No. 0.00071		
Matches 436	Conservative	0	Mismatches 576	Indels 3
				Gaps 1
QY	84	lgaagcggtacagatgtatgcacacgttagtaagccaaagcaaacgacatgccaataaacg	143	
Db	915	TGAAGAGAAAAAGAGAGAGTACACATGAGAGAAATACGAAAAAGAGAGCATGAGAGAGT	974	
QY	144	atactactcttaacagccctacagtaacggaacatggtgtaattcccaacattacgattc	203	
Db	975	AATACATGAGAGAGAAAAAGAGAGAGTACACATCAAGAAAAAGAAAAAGAGAGACATCA	1034	
QY	204	atalgtcgtatatacaacaagacgaaaaaacgctaccctggtgtgcgttagcatgtatgataa	263	
Db	1035	AGAGATATATACATGAGAGAGAAAAAGAGAGAGTACACATCAAGAGAGAGAGAGAGAGT	1094	
QY	264	agcaggtgtgctgcgaaaaaagacgcttacttagcttgatttaccaaaaagaatalgaactta	323	
Db	1095	AACACATGAGAGAGAGAAAAAGAGAGTATCATGATCAAGAGAGAAAAAGAGAGAGTATCA	1154	
QY	324	cgtttccaagcaaaccttaatactctgcggaagctcggttagaacttaccatgcattgctta	383	
Db	1155	TGAGAGAGAGAAAAAGAGAGAGTACCATGAGAGAGAGAAAAAGAGAGAGTACCATGAGAG	1214	
QY	384	caactatgcaacaaatattagacgaaatgctccaaagagctagaagctctgttcaagcaaa	443	

Db	1215	AGAAAAAGACAGTAGTAACATGAAAGAAAGAAAAAGACAGTAGTAACATGAAAGAAAGAAA	1274
Qy	444	agatttaaanaagagaacaatactatccacanaaatttcctatagaattaanaactcgcac	503
Db	1275	ACAAGACGTAAACATGAAAGAAAGAAAGTAACATGAAAGAAAGAAAGAAAGAGT	1334
Qy	504	agtcattttagatcgcgctataltglaaacaactcgtgatttacttcgcttcaattaa	563
Db	1335	AACACATGAGAGAGAAAGAAAGAAAGTAACATGAGAGAGAGAAAGTAACATGTA	1394
Qy	564	agcaaaagcacagaacttcgagcagcttaatttaigtatattaccgttgaataagaagc	623
Db	1395	ACAAGACAGAAAGTAACATGAAAGAAAGAAAGTAACATGTAAGAAAGAAAGAA	1454
Qy	624	gagcgaaatcacagacgcgtgtgaagcagcgcaatttaagacaagaagtacgtctgtga	683
Db	1455	ACAAGACAGAAAGTAACATGAAAGAAAGTAACATGAAAGAAAGTAACATGAAAGAA	1514
Qy	684	tcaattcaatcaatcttaccaaaagtacagatgcttccaacttcaacagaagt	743
Db	1515	AGTAACATGAGAGAGAAAGAAAGTAACATG---AGAGAGAAAGAAAGAAAGAA	1571
Qy	744	agcgaaaaaagatttagatgcagatgaagcgtcgcttactccaagaattgaagtgtag	803
Db	1572	TCAGGACAGATGAGGA	1631
Qy	804	tgcgattcaactcaaaaacaagctgcttgaatttaacagcaatgacagtgaaacgcacact	863
Db	1632	AG	1691
Qy	864	aaaattcaacttccagctgctcgaatgaatgaatgaatgaatgaatgaatgaatgaat	923
Db	1692	AG	1751
Qy	924	ctataaagtgcagcgttaacattccatttgccttataatgcgagtgctttctttctac	983
Db	1752	AACACATGAG	1811
Qy	984	agacggaanaaacatcactcgtgtgactgcttcaactccattcggaaaataaatacgaagtata	1043
Db	1812	AG	1871
Qy	1044	agtagtagttaaggtatttaagacaaaatgycagaagaatttaagaagaatgca	1098
Db	1872	AAAGTAGACATGAG	1926
RESULT	8		
EI0126			
ID	EI0126	standard; DNA; UNC; 3399 BP.	
XX	XX	EI0126:	
SV	SV	EI0126.1	
DT	DT	08-OCT-1997 (Rel. 52, Created)	
DT	DT	02-SEP-2000 (Rel. 65, Last updated, Version 2)	
DE	DE	DNA encoding an immunogenicity protein of leucocytozoon caulleryi fused to maltose-binding protein.	
DE	DE		
XX	XX	JP 1995284392-A/2.	
OS	OS	unidentified	
OC	OC	unclassified.	
XX	XX		
XX	XX	[1]	
RA	RA	Doi H., Nagakuchi Y., Tanaka Y., Fujisaki Y.;	
RT	RT	"GENE CLONE MANIFESTING CHICKEN LEUCOCYTOZOON IMMUNOGENIC PROTEIN AND GENE RECOMBINANT VACCINE AGAINST CHICKEN LEUCOCYTOZOON";	
RL	RL	Patent number jp1995284392-A/2, 31-OCT-1995.	
RL	RL	DOUBUNYOUN SEIBUTSUGAKUTERI SEIZAI KYOKAI, KITASATO INST:THE.	


```

Oy 131 agtcaaaaagcatactactactactacagccatcacglaacggaacgtgtgaattcccaa 190
Db 31458 AAGAAAGCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 31399
Oy 191 acattaacgattatattgtcgtatcaacaagaacgaaaaacgataccgtgacgttag 250
Db 31398 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 31339
Oy 251 catttgtaataaagcaggtggtgcgaaaaaagcgttactactgtgtttacaaaag 310
Db 31338 CAGAAGCAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGCAG 31279
Oy 311 aatgaagaacttacgttttcaaaagcaaaccttaactctgcgaagctgcgtgacactt 370
Db 31278 CAGCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 31219
Oy 371 acatcgatgcttacaactatgcacaanaattagacgaaatgcgcgaagagctagagctg 430
Db 31218 AAGAAGCAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAG 31159
Oy 431 cgttcaagcaaaagatttagaaaaagcgaacaatactactacaaa 476
Db 31158 AAGAAGCAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAGCAAGAAAG 31113

```

RESULT 13

PFMAL4P1.1
WPCOMMENT
Sequence split into 4 fragments LOCUS PFMAL4P1 Accession AL034557

Fragment Name Begin End
PFMAL4P1_0 1 110000
PFMAL4P1.1 100001 210000
PFMAL4P1.2 200001 310000
PFMAL4P1.3 300001 392633
Continuation (2 of 4) of PFMAL4P1 from base 100001 (AL034557 Plasmidium falciparum chrom

Query Match 1.7%; Score 64.2; DB 84; Length 110000;
Best Local Similarity 43.2%; Pred. No. 0.081;
Matches 306; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

```

Oy 197 acgatgatatatgctgaatacaacaagaacgaaaaacgataccgtgtacgttagcattg 256
Db 87711 ATGAAGAGCATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGATG 87770
Oy 257 tgaataaagcaggtggtgcgaaaaaagcgttactactactgattacaaaagaatatg 316
Db 87771 AAGATGAGCAAGATGAAGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 87830
Oy 317 aaacttaacgtttcaaaagcaaaccttaactctgcgaagctgcgttagcaacttacatcg 376
Db 87831 GAATGACAGATGGCGAAGAAATGATGATTAATGAAGAAAGATGATGATGATGATGATG 87890
Oy 377 atgcttaacactatgcaaaaaaattagacgaaatgcgcgaagcttagaggtctgtcttc 436
Db 87891 ATGATGAAGAAAGAAATGATGATGAAGAAAGAAATGATGATGATGATGATGATGATG 87950
Oy 437 aagcaaaagattagaaaaagcagaacaatactactatcaaaaacttcttaagaaatta 496
Db 87951 TGAAGAAATGATATACGAAATGATGATGACGAAAGAAATATGATTAACGAAATGATG 88010
Oy 497 ctgcagacgtacatttagatcgcgatatagtgtaaaaacaactcgtgattacttcgtcta 556
Db 88011 ATGACGTGAAGAAATGATGATGACGTGAAGAAATGATGATGATGATGATGATGATGATG 88070
Oy 557 catttaaaagcaaaacacaaagactcgcgcagacttaattatgatatattacgtttgcaa 616
Db 88071 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 88130
Oy 617 tgaagcgcgcgaagctacaaagcgtgtgtgaaagcggcgaatttagacaagactaaagctg 676
Db 88131 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 88190

```

```

Oy 677 cgttgatcaaatcaatcaactactaccaaaagtaacagatgcttcaaaactgaactaa 736
Db 88191 ATGATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 88250
Oy 737 cagaagtagcgaanaagatcagatgcagatgaagctgcgttactcctcaaaagtggaa 796
Db 88251 AAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 88310
Oy 797 gtgtagtgcgatttaacactcaaaaagcgtgtgtaataacagcagttaccagtggaac 856
Db 88311 ATGATTAATGACGATGTTAATGAATTAATGAAGAAATTAATTAATTAATTAATTA 88370
Oy 857 gacaactaaattacaacttcagctgcgtgcacgaatcagtaaa 905
Db 88371 CTTTACATTTAAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 88419

```

RESULT 14

S76368/c
LOCUS
DEFINITION
ORF 5' of ECUF2; ECRF3-C protein-coupled receptor homolog [herpesvirus Saimiri] HVS, host-squirrel monkey, Genomic, 4 genes, 3720 nt].

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Saimirine herpesvirus 2 host-squirrel monkey.
Saimirine herpesvirus 2
Viruses; dsDNA viruses, no RNA stage: Herpesviridae;
Gammaherpesvirinae; Rhadinovirus.
1 (bases 1 to 3720)
Nicholas,J., Cameron,K.R. and Honess,R.W.
Herpesvirus saimiri encodes homologues of G protein-coupled
receptors and cyclins
Nature 355 (6358), 362-365 (1992)
92115001
GenBank staff at the National Library of Medicine created this
entry [NCBI g1bseq 76368] from the original journal article.
This sequence comes from fig1b.
Location/Qualifiers
1..3720
/organism="Saimirine herpesvirus 2"
/db_xref="taxon:10381"
complement(1..194)
/partial
/gene="ORF 5' of ECUF2"
complement(1..183)
/partial
/note="ORF 5' of ECUF2"
/note="This sequence comes from fig1b"
/codon_start=1
/protein_id="AAB21118.1"
/db_xref="GI:243352"
/translation="MDLKTIVLHITDSEEMKCLFLINGCIPRSCNAVNSIDLIIE
ETLSKSTQMDICLMQCL"
complement(95..977)
/gene="ECUF2"
/note="cyclin homolog"
complement(184..948)
/gene="ECUF2"
/note="cyclin homolog: This sequence comes from fig1b"
/codon_start=1
/protein_id="AAB21115.1"
/db_xref="GI:243352"
/translation="MADSPNRLNRAKIDSTTKMDPRVNLNKKLELLIPKTSLSMEIQ
TEVIVDNRTLLTMMHLLCESFELDKSVFPLSVILDRYLCKRGKTKTLOKIAACV
LIGSKIRTVKPMVSKLYLSCDFTNLELINOXKIDLEALKMDTEAVLATDFLIPIC
NALKIPEDLPOLYEASTTICKALIDPNTALISPGILICAGGLITTTIETDNTNRPMT
CYLBDLSILNESTNVTYRVKDYSEAFSLYDEIL"
complement(978..2476)
/gene="ORF 5' of ECRF3"
complement(1004..2227)

gene
CDS

THIS PAGE BLANK (USPTO)